Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
Air Force Personnel and Training Research Center

DAVID A. GRANT, Associate Editor
University of Wisconsin

DELOS D. WICKENS, Associate Editor
Ohio State University

Consulting Editors

E. JAMES ARCHER, *University of Wisconsin*
JUDSON S. BROWN, *State University of Iowa*
CLETUS J. BURKE, *Indiana University*
PAUL M. FITTS, *Ohio State University*
FREDERICK C. FRICK, *Massachusetts Institute of Technology*
FRANK A. GELDARD, *University of Virginia*
JAMES J. GIBSON, *Cornell University*
CLARENCE H. GRAHAM, *Columbia University*
HAROLD W. HAKE, *University of Illinois*
LLOYD G. HUMPHREYS, *AF Personnel and Training Research Center*
ARTHUR L. IRION, *Tulane University*
HOWARD H. KENDLER, *New York University*
DONALD B. LINDSLEY, *University of California, Los Angeles*
KENNETH MACCORQUODALE, *University of Minnesota*
QUINN MCNEMAR, *Stanford University*
NEAL E. MILLER, *Yale University*
KENNETH W. SPENCE, *State University of Iowa*

ARTHUR C. HOFFMAN, Managing Editor
HELEN ORR, Assistant Managing Editor
Editorial Staff: SADIE J. DOYLE, BARBARA CUMMINGS, FRANCES HAZEL

CONTENTS

Physical Determinants of the Judged Complexity of Shapes: F. ATTNEAVE 221

Avoidance Learning Motivated by Hypothalamic Stimulation: B. D. COHEN, G. W. BROWN, AND M. L. BROWN 228

An Analysis and Revision of Hull's Theorem 30: A. C. PEREBOOM 234

Certain Relationships Between Stimulus Intensity and Stimulus Generalization: W. HEYMAN 239

The Oculomotoric Pattern of Circular Eye Movements During Increasing Speed of Rotation: S. J. GERATHEWOHL, H. STRUGHOLD, AND W. F. TAYLOR 249

Discrimination Learning and Transposition in Children as a Function of the Nature of the Reward: G. TERRELL, JR., AND W. A. KENNEDY 257

Effect of Control Lag on Performance in a Tracking Task: J. E. CONKLIN 261

Nonreinforcements Versus Reinforcements as Variables in the Partial Reinforcement Effect: H. H. KENDLER, S. S. PLISKOFF, M. R. D'AMATO, AND S. KATZ 269

Effects of Foreperiod, Induced Muscular Tension, and Stimulus Regularity on Simple Reaction Time: W. H. TEICHNER 277

American Psychological Association

Vol. 53 No. 4 April 1957

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $8.00, or $16.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This JOURNAL publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects, except when the latter are specifically oriented toward the extension of behavior theory. Experimental psychometric studies and studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology.

Normally, articles exceeding 20 printed pages cannot be accepted. Within this limit, the piecemeal, experiment-by-experiment reporting of psychological research is discouraged, and the reporting of the data and conclusions of substantial segments of research effort is encouraged. Specifically, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) *must* be presented in a single article. Special provision is made for the publication of Supplementary Reports and Replication Reports in articles of not more than 1.5 printed pages (see Editor's Note, this JOURNAL, 1957, 53, 1–2).

Address all articles submitted for publication to the editor, Arthur W. Melton, Air Force Personnel and Training Research Center, P. O. Box 1557, Lackland Air Force Base, Texas. All manuscripts should be submitted in two complete copies; only the original copy will be returned if the article is not accepted for publication. Figures may be original line drawings or glossy photographic prints, but should not exceed 8½ by 11 in. Care should be taken to insure legibility of labels and critical features of drawings when reduced to a single-column or (exceptionally) a double-column figure in the JOURNAL.

To save printing costs supplementary material to articles in this JOURNAL are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

Manuscripts submitted to the JOURNAL *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc. as described in the Publication Manual of the American Psychological Association (*Psychol. Bull.*, 1952, 49 [4, Suppl.], 389–449). In particular, the organization of the manuscript should be text, references, footnotes, tables (one to a page), figure titles (typed on separate page), and figures. All typed material *must* be double-spaced with ample margins. The term "subject," with its various forms, is abbreviated *S*, *Ss*, *S*'s, and *Ss*'; analogous abbreviations are used for "experimenter" (*E*) and "observer" (*O*). When in doubt about the practices of the JOURNAL, authors should examine a recent issue.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879. Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1957 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 53, NO. 4 APRIL, 1957

PHYSICAL DETERMINANTS OF THE JUDGED COMPLEXITY OF SHAPES[1]

FRED ATTNEAVE

Operator Laboratory, Air Force Personnel and Training Research Center

Many psychological tasks vary in difficulty with the complexity of the stimulus objects involved. Complex visual objects are not only harder to reproduce from memory than simple ones (**4, 6**), but also harder to learn by name (**4, 7**) and to match (**11**). *Complexity* is an ill-defined variable, however. No two of the experiments referenced above employ exactly the same operations of physical measurement, and an essential communality between them is not easy to specify in objective terms.

In the study reported here, ratings of the complexity of nonrepresentational shapes were obtained from a large number of *S*s and related to measurable physical characteristics of the shapes. The results not only have interest in their own right, but also serve to indicate the physical variables most likely to be relevant to other tasks, like those initially mentioned, on which data are typically harder to get and less precise. The relationship of judged complexity to informational content, "degrees of freedom," compactness, and certain other variables will be considered.

[1] This research was carried out at the Skill Components Research Laboratory, Air Force Personnel and Training Research Center, Lackland Air Force Base, San Antonio, Texas, in support of Project 7706. Permission is granted for reproduction, translation, publication, use, and disposal in whole or in part by or for the United States Government.

METHOD

Materials.—The stimuli were 72 "random" shapes, each constructed by the following general method. In a $k \times k$ matrix, n random points were plotted, i.e., all coordinates of the points were random numbers between 1 and k. The points were then connected randomly into a polygon of n sides. This connecting process involved two steps. First, peripheral points were connected into a convex polygon which enclosed all the points not included in its contour (as if a pin were stuck into each point and a rubber band snapped around the whole cluster). Second, the unconnected points were given a random order and each in turn was "taken into" a randomly chosen segment of the surrounding polygon (as if by hooking that segment of the rubber band over the interior pin). Since lines connecting the points were not permitted to cross, the number of alternative segments into which a given point might be taken could either increase or decrease as the process continued; hence the assignment of a random sequence to the unconnected points. Connections which placed certain points outside the polygon were permitted, though in consequence the ways in which such points could themselves be connected were restricted.

Some of the polygons thus constructed were further developed into curved figures by replacing angles with inscribed arcs of random

curvature. If the shorter of the two segments forming an angle was between d (an integer) and $d+1$ matrix units in length, a random number between O and d was taken as the distance from the vertex of the angle to the point of tangency of the arc to be inscribed. [An illustrated and more detailed description of the construction methods outlined above is presented in a methodological paper by Attneave and Arnoult (5).]

A class of symmetrical figures was also constructed by the following operations. An asymmetrical polygon was reflected about a vertical axis passed through the point farthest to the right in the original polygon. This reflection resulted in a figure consisting of two parts: the original and its mirror-image, touching (ordinarily) at a single point—the point of reflection. In order to tie the two halves together into a single unified shape, a part of the area between them was filled in—specifically, the area between that pair of touching symmetrical segments which had the greater vertical component. Reflecting a shape resulted in a corresponding increase in area.

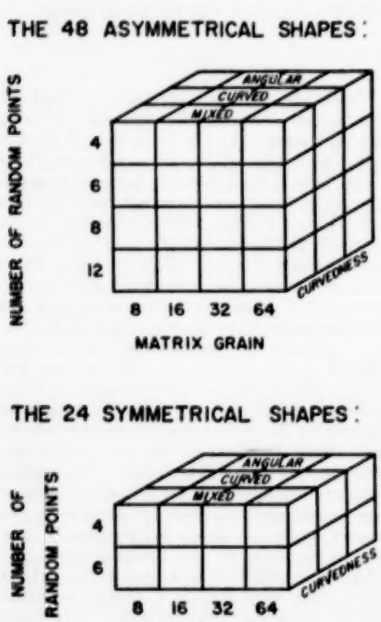

FIG. 1. Plan by which the shapes were systematically varied.

In the construction of curved symmetrical shapes, the figure was first reflected and then curved—all curves being repeated symmetrically—except in the unique case of the arc associated with the point of reflection.

The following parameters of the shapes were varied in a quasi-factorial design (see Fig. 1):

1. *Matrix grain.* Four different matrix grains (8×8, 16×16, 32×32, and 64×64) were used in plotting the random points from which the shapes were constructed. Eighteen shapes were made with each grain. This variation in grain did not involve a concomitant variation in size. Moreover, the matrix lines did not appear when the shapes were displayed to *Ss*. The amount of information required to locate each point in the matrix (i.e., the amount of information taken from the random number table) was very nearly equal to 6, 8, 10, or 12 bits, depending upon the fineness of grain (these values are a trifle too great, since the selection of each point reduced by one the number of alternative positions which the next point might take).

2. *Curvedness.* One-third of the shapes, or 24, were entirely angular, one-third were entirely curved (i.e., all angles were replaced with inscribed arcs), and the remaining third were mixed. For each mixed shape, both the number and the identity of angles to be replaced with curves were randomly determined.

3. *Symmetry.* Two thirds of the shapes, or 48, were asymmetrical; the remaining third were symmetrical. The reason for this unequal division will be discussed in connection with the next variable.

4. *Number of turns.* Either 4, 6, 8, or 12 random points were initially plotted for the construction of each shape. The asymmetrical shapes were evenly distributed over these four classes with 12 in each class. The 24 symmetrical shapes, however, were equally divided between the 4- and 6-point classes. They were thus equivalent to the 4- and 6-point asymmetrical shapes with respect to number of *independent* sides or contour turns, and approximately equivalent to the 8- and 12-point asymmetrical shapes with respect to *total* number of sides or turns. (In referring to the present variable, we shall hereafter use the term *turns* in preference to "points," "angles," or "sides," since it applies equally well to curved and angular figures.) In the process of joining points into polygons, a point was occasionally "lost" when the two lines meeting at the point themselves formed a straight line; hence it was possible for a 12-point figure, for example, to have only 11 or even fewer turns in its contour. The actual number of turns, rather than the original number of points, was considered to be the

important variable in the subsequent analysis of data.

The four variables just enumerated were orthogonally related, as indicated in Fig. 1, and each possible combination of values on these variables was represented in one and only one of the 72 shapes, except as otherwise noted in connection with Symmetry and Number of Turns. Six of the shapes are reproduced in Fig. 2 to give the reader a general idea of their appearance. It should be understood that every shape was constructed from a completely independent selection of random points; e.g., the same shape never appeared in an angular and a curved version, nor in an asymmetrical and a symmetrical version.

In addition to these systematically varied parameters, the following two physical measures were taken after the shapes were constructed:

5. P^2/A. The square of perimeter, divided by area, was obtained for each shape. This is a size-invariant measure of dispersion, or non-compactness; see Attneave and Arnoult (**5**).

6. *Angular variability.* This measure represents the average *difference* between adjacent angles in a polygon. More precisely, it is the arithmetic mean (sign ignored) of the algebraic differences in degrees of slope-change (sign observed) between all successive or adjacent angles taken in overlapping pairs about the contour, convex angles being considered positive and concave ones negative. For example: suppose that the slope-changes associated with the angles of a four-sided figure are 150, 130, − 90 (a concave angle), and 170 degrees. The differences between adjacent angles will be 20, 220, 260, and 20 degrees. Angular Variability is the mean of these four differences, or 130. In the case of curved figures, measurement was made upon the polygons from which the curved shapes were constructed. Something like "good continuation" is quantified in this measure; a value of zero means that all angles of the polygon are equal, whereas high values are associated with jagged, irregular figures. As one might expect, this is correlated with the preceding measure, P^2/A: $r = .48$ over the 72 shapes.

Subjects.—One hundred and sixty-eight airmen basic trainees at Lackland Air Force Base rated the shapes for complexity.

Procedure.—The shapes were displayed, in a white-on-black figure-ground relationship, by projecting them individually to the front wall of the room in which *S*s were seated. Each *S* was provided with a seven-category rating form containing a line for each shape and columns headed "Extremely Simple," "Very Simple," "Simple," "Medium," "Complex," "Very Complex," and "Extremely Complex." The instructions given contained no definition, either explicit or implicit, of the terms "Simple" and "Complex."

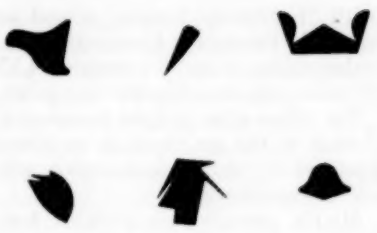

FIG. 2. Six of the 72 shapes used in the study.

Before actually making their ratings, the *S*s were shown all 72 shapes in rapid succession (2-sec. exposure each) in order that they might, from the beginning, adapt their rating behavior to the range of stimuli with which they were to deal. In displaying the shapes for rating, each was exposed for 10 sec. with a negligible interval between stimuli.

The *S*s served in four approximately equal groups which differed only in the ordering or sequence of stimulus presentation. The first sequence was a random permutation of the 72 stimuli, the second was an "inside-out" reordering of the first, and the third and fourth were reversals of the first two.

RESULTS

Ratings were scaled by the method of graded dichotomies (**1**)[2] in order to render them comparable to paired-comparison scale values. It turned out that mean ratings might as well have been used, since a plot of means-versus-scaled ratings showed little or none of the usual curvilinearity. The linearity of this relationship implies that the "widths" of the various scale categories were very nearly equal in terms of the dispersion of response distributions, though *S*s had no special instructions to regard them as equal. Mean ratings ranged from .82 to 5.24

[2] This method, or the method of successive intervals, has been independently discovered and rediscovered, with minor variations, by Urban, Guilford, Thurstone, and me [see Guilford (**8**, Ch. 10)]. I continue to use my own version simply because it is the one most familiar to me.

(with "Extremely Simple" scored as zero and "Extremely Complex" as 6). Scaled ratings covered a range of 4.17 *SD* units with an arbitrary zero point.

The effect upon judged complexity of each of the six physical variables described in the previous section will now be considered.

Matrix grain.—This variable had no measurable effect on the ratings. In the original data there was a slight, nonsignificant suggestion of a trend toward greater complexity for finer grains, but even this failed to appear in the residuals from subsequent variables (*vide infra*).

Curvedness.—Likewise, whether the shapes were angular, curved, or mixed made no significant difference in judged complexity. After the removal of variance attributable to the important variables Symmetry, Number of Turns, and Angular Variability, a simple analysis of variance for effects of Curvedness was performed on the residuals. This yielded an F of 1.73 with $df = 2/69$, which is far short of the 3.13 required for the 5% level of significance. The mean rating obtained for curved shapes was lower by about .15 scale unit (i.e., *SD* units) than the means for angular and mixed shapes, which were almost exactly equal. Even if this difference were reliable, it would account for less than half of one per cent of the total variance of the ratings.

Symmetry.—Symmetrical shapes were judged *more* complex than asymmetrical shapes with the same number of *independent* turns by a mean of about .41 unit ($P = .001$, $t = 3.99$, $df = 69$). They were judged *less* complex than asymmetrical shapes with the same *total* number of turns by a somewhat greater amount, i.e., about .9 or 1.0 units. Reflecting a shape symmetrically is equivalent to an increase of about 19% in number of independent turns, in terms of effect on judged complexity (see Equation 1, below).

Number of turns.—This was by far the most important variable. Plots of rated complexity vs. number of independent turns showed appreciable curvilinearity, which was rectified when the logarithm of the number of independent turns was used instead. Regression lines relating complexity to log turns were found, by the least-squares method, for symmetrical and asymmetrical shapes separately. Associated with the former was a correlation coefficient of .75; with the latter, a coefficient of .93. This difference is attributable chiefly to the restricted range of the symmetrical shapes. It was possible to adjust the lines to equality of slope, by the use of a weighted average, with negligible impairment of goodness of fit; the resulting parallel lines were separated by the .41 units mentioned above.

The relationships thus far were combined into the following equation:

$$J = 5.46 \log_{10} T + .41 S - 2.30 \quad (1)$$

in which J is judged complexity (on a scale in which the simplest figure is arbitrarily assigned a value of zero), T is number of independent turns, and S is a variable on which symmetry has a value of 1 and asymmetry a value of 0. This equation accounts for 82.5% of the total variance of the complexity ratings (see Table 1). For purposes of further analysis, a table of actual residuals from Equation 1 was made up by subtracting the predicted rating of each shape from the obtained rating.

P^2/A and Angular Variability.—It will be recalled that these two physical measures were intercorrelated to the extent of .48. Each was now correlated with the residuals from Equation 1. Angular Variability showed

TABLE 1

CONTRIBUTIONS OF THE INDIVIDUAL PHYSICAL VARIABLES TO JUDGED COMPLEXITY

Physical Variable	Percentage Variance of Complexity Ratings Explained
Matrix Grain	...
Curvedness	...
Symmetry	3.8
Number of Turns	78.7
P^2/A	0.9*
Angular Variability	7.1
Total	90.5

* P^2/A actually explains 4.5% of the variance of the ratings, but of this 3.6% is shared with Angular Variability. The shared 3.6% is arbitrarily included in the 7.1% attributed to Angular Variability.

the higher relationship, $r = .64$, as compared with $r = .51$ for P^2/A. One may question whether P^2/A accounts for any variance not also accounted for by Angular Variability. The answer is probably affirmative, since the partial correlation between P^2/A and the residuals, with Angular Variability held constant, is equal to .30, which is on the borderline of significance at the 5% level. Bearing in mind that the residuals from Equation 1 themselves represent only 17.5% of the total variance of the ratings, we may calculate that 3.5% of the total variance is uniquely predicted by Angular Variability, that .9% is uniquely predicted by P^2/A, and that a further 3.6% is predictable from either Angular Variability or P^2/A, or common to the two predictor variables.

The relationship between Angular Variability and the residuals from Equation 1 was also considered for symmetrical and asymmetrical shapes separately. The correlation coefficient was .77 for the former and .57 for the latter. After equating the regression lines in slope, the line for asymmetrical shapes was about .13 unit higher than that for symmetrical shapes (this constant is accordingly added to the coefficient of S in Equation 2 below). However, the gain in predictive efficiency which results from treating symmetrical and asymmetrical shapes separately at this stage is scarcely worth bothering about; it amounts to only about .6% of the total variance of the ratings.

Adding Angular Variability as a predictor variable:

$$J = 5.46 \log_{10} T + .54\, S + .005\, AV - 2.91. \quad (2)$$

A second table of residuals from this equation was prepared and examined for possible minute effects of Matrix Grain and Curvedness, with the negative results mentioned earlier. It may be noted that although Curvedness per se has no demonstrable effect on the ratings, Curvedness is nevertheless represented in the variable P^2/A (i.e., the curving of any shape by the method used necessarily decreases its P^2/A value), and the component of P^2/A which is affected by Curvedness does attain a marginal level of significance (see above).

DISCUSSION

The lack of effect of Matrix Grain on the judgments, though not intuitively surprising, does serve to demonstrate how the amount of information contained in a stimulus (from the experimenter's point of view) may be varied greatly without changing the apparent complexity of the stimulus. The amount of information gained by S in viewing a shape may very well be independent, within broad limits, of the amount of information used in this way by E in constructing the shape.

More interesting is the discovery that curved shapes are judged no more complex than angular shapes. One might reasonably expect that apparent complexity would depend upon the number of dimensions or degrees of freedom

associated with a stimulus, i.e., upon the number of numbers necessary to describe it in some coordinate system [this is essentially what MacKay (9) has called "logon content," and Pollack and Klemmer (10) have called "coordinality"]. This hypothesis correctly predicts the effect of Number of Turns upon judged complexity, since the number of coordinates per turn is a constant for either angular or curved shapes, considered separately. But whereas only two numbers are necessary to describe each turn in an angular shape (i.e., the x and y coordinates of the turning point), an additional number (i.e., radius of curvature) is necessary at each turn in a curved shape. Thus, curved shapes not only contain more information, but also require more dimensions for their specification; nevertheless they appear no more complex than angular ones. Another hypothesis which would predict higher values for the curved shapes is that complexity is a function of the total number of individually homogeneous parts, e.g., lines of constant slope, or curves of constant curvature, in the contour [cf. the measure of complexity used by Fehrer (6)].

The finding that symmetrical figures are intermediate between asymmetrical figures with the same number of independent turns and those with the same total number of turns is consistent with results previously obtained in a group of learning studies (4) in which dot patterns were used as stimuli.

The Angular Variability measure, no less than Symmetry, has to do with repetitiveness or similarity between parts of a contour; e.g., to state the limiting case, any polygon which is regular in the sense that all angles are equal will have a value of zero on this variable.

The Number of Turns variable, which accounted for nearly four-fifths of the variance of the judgments, may be described as the number of maxima (regardless of sign) in one cycle of the function relating curvature to distance along the contour. This function is a series of spikes for any angular shape, and a step-function for any curved shape constructed by the present method (5). In the case of the latter function, information is concentrated in the vertical transition lines; i.e., in points on the contour at which arcs are tangent to straight lines. This is true from the point of view of a hypothetical observer who is in some sense capable of extrapolating curvature. From the point of view of an observer incapable of extrapolating curvature, however, but capable of extrapolating slope, information is associated simply with the maxima of the curvature function; i.e., information, or uncertainty, is concentrated in the arcs of the contour, but uniformly distributed within any given arc (though by some "averaging" process the observer might associate the information more specifically with the midpoint of the arc).

In an earlier study (2), *S*s tended to choose maxima of curvature when asked to select the points most representative of a given contour. It may be that an observer's primary impression of complexity depends upon the number of such maxima in a shape and that this first estimate is then reduced more or less by the presence of certain types of redundancy or repetitiveness to which he is sensitive. The possibility of devising a single physical measure which would predict apparent complexity as efficiently as the combination of predictor variables which we have considered is not to be dismissed, but appears unlikely.

Summary

Judgments of the complexity of 72 shapes were obtained from 168 *S*s. The shapes were constructed by a method in which certain physical characteristics were systematically varied and the remainder randomly determined. About 90% of the variance of ratings was explained by (*a*) the number of independent turns (angles or curves) in the contour, (*b*) symmetry (symmetrical shapes were judged more complex than asymmetrical with number of *independent* turns constant, but less complex with *total* number of turns constant), and (*c*) the arithmetic mean of algebraic differences, in degrees, between successive turns in the contour. Angular and curved shapes were judged about equally complex, though the latter involved additional degrees of freedom (radii of curvature). Also

immaterial, within broad limits, was the grain of the matrix from which critical points were chosen to construct the shapes.

References

1. Attneave, F. A method of graded dichotomies for the scaling of judgments. *Psychol. Rev.*, 1949, **56**, 334–340.
2. Attneave, F. *The relative importance of the parts of a contour.* San Antonio, Tex.: Human Resources Research Center, Lackland Air Force Base, November 1951. (*Res. Note* P&MS 51-8.)
3. Attneave, F. Some informational aspects of visual perception. *Psychol. Rev.*, 1954, **61**, 183–193. (Distributed separately as *Res. Rep.* AFPTRC-TR-54-33, August 1954).
4. Attneave, F. Symmetry, information, and memory for patterns. *Amer. J. Psychol.*, 1955, **68**, 209–222. (Distributed separately as *Res. Rep.* AFPTRC-TN-55-29, October 1955).
5. Attneave, F., & Arnoult, M. D. The quantitative study of shape and pattern perception. *Psychol. Bull.*, 1956, **53**, 452–471.
6. Fehrer, E. V. An investigation of the learning of visually perceived forms. *Amer. J. Psychol.*, 1935, **47**, 187–221.
7. French, R. S. Identification of dot patterns from memory as a function of complexity. *J. exp. Psychol.*, 1954, **47**, 22–26.
8. Guilford, J. P. *Psychometric methods.* New York: McGraw-Hill, 1954.
9. MacKay, D. M. The nomenclature of information theory. In H. von Foerster (Ed.), (*Transactions of the eighth conference on*) *Cybernetics; circular causal and feedback mechanisms in biological and social systems.* New York: Josiah Macy, Jr. Foundation, 1952. Pp. 222–235.
10. Pollack, I., & Klemmer, E. T. *The assimilation of visual information from linear dot patterns.* Cambridge, Mass.: Air Force Cambridge Research Center, July 1954. *Tech. Rep.* (AFCRC-TR-54-16.)
11. Weinstein, M. Stimulus complexity and the recognition of visual patterns. Unpublished doctor's dissertation, The Ohio State Univer., 1955.

(Received March 16, 1956)

Journal of Experimental Psychology
Vol. 53, No. 4, 1957

AVOIDANCE LEARNING MOTIVATED BY HYPOTHALAMIC STIMULATION[1]

BERTRAM D. COHEN,[2] GEORGE W. BROWN,
Veterans Administration Hospital, Iowa City, Iowa,
AND MARJORIE L. BROWN
The State University of Iowa

When an electrical stimulus is delivered to the hypothalamus of a cat, a characteristic syndrome of manifest emotionality is produced, e.g., hissing, spitting, pilo-erection, pupillary dilation, etc. (**1, 9**). This syndrome has been variously termed "sham rage," "pseudo-affective," and "quasi-emotional." An instrumental conditioning experiment reported by Masserman (**5**) was performed to test the hypothesis "that direct hypothalamic stimulation produced not only the external manifestations of anger, but also a true affective experience which would be anticipated at the sensory signal and for which the animal could learn to prepare, compensate, or adapt" (**5,** p. 11). Briefly, Masserman attempted to train a group of cats to avoid hypothalamic stimulation by responding adaptively to a set of lights, sounds, and airblasts which were regularly presented just prior to the hypothalamic stimulus. A control group was given the same training procedure with the exception that pain stimulation produced by a grid shock was administered in place of the direct hypothalamic stimulation. The control *Ss* were found capable of responding appropriately to the sensory signals. They avoided the grid shocks "either by mounting a small platform in the cage or by hanging from a bar near the roof." The sensory signals apparently acquired no such meaning for the experimental *Ss*; they learned no equivalent instrumental response.

Masserman concluded that, in contrast to a grid shock, direct hypothalamic stimulation cannot provide the motivation necessary for the learning of an instrumental avoidance response. This conclusion is questionable, however, in view of the inability of *Ss* in the experimental group to select and perform a specific "escape" response under hypothalamic arousal conditions. The escape response in the control procedure was determined by the spatial localizability of the pain-producing grid. That is, the adaptive (pain terminating) escape response consisted in *Ss* doing something which put distance between themselves and the grid. From what were they to escape when excited by the central hypothalamic stimulus? It is doubtful that, without special pretraining or selection, cats would regularly jump to a small platform or bar under "sham rage" conditions. If they could not perform an escape response to hypothalamic stimulation, one could not reasonably expect them to perform in an anticipatory avoidance fashion.

In order to provide *Ss* with a convenient adaptive response to centrally produced emotional excitement, Delgado, Roberts, and Miller (**4**)

[1] This investigation was supported in part by a research grant M-641-C from the National Institute of Mental Health, of the National Institutes of Health, U. S. Public Health Service.

[2] Now at Lafayette Clinic, Detroit, Michigan.

first pretrained cats to terminate a painful grid shock by turning a wheel. The *S*s were also trained to avoid the grid shock by turning the wheel in response to a buzzer which regularly preceded the grid shock. After this pretraining, the wheel-turning response was found capable of evocation by subcortical stimulation. Following this, the central stimulus was appropriately paired with neutral sensory stimuli. The pretrained adaptive response was, in this manner, successfully conditioned to the sensory stimuli.

These results suggest that the emotional response components produced by the central stimulation overlapped those produced by the peripheral pain (grid shock) and threat-of-pain (buzzer) stimuli. The internal cues produced by the central stimulation could therefore evoke the same adaptive response previously associated with pain and threat of pain. In addition, the motivational properties of the centrally produced emotionality could provide a basis for conditioning the pretrained habit to new sensory stimuli.

The pretraining technique employed by Delgado *et al.* provided *S*s with an available instrumental response. Masserman's negative findings may have been obtained because no equivalent adaptive response was included in the behavioral repertoires of his *S*s under hypothalamic arousal conditions.

Another difference between these two studies is in the brain regions stimulated. Delgado *et al.* placed electrodes in the tectal area, the lateral nuclear mass of the thalamus, and the hippocampus. Masserman's report concerned only the effects of hypothalamic placements.

The purpose of this paper is to present evidence that direct hypothalamic stimulation can provide the motivation necessary (*a*) for the production of instrumental escape behavior, (*b*) for the conditioning of an emotional response to a neutral external stimulus, and (*c*) for the maintenance of an avoidance response based on the learned motivation provided by (*b*).

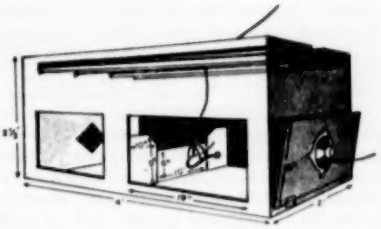

FIG. 1. The conditioning apparatus.

METHOD

A Labtronics Stereotaxic apparatus was used to implant permanent electrodes in the hypothalamus of four cats. The electrodes were made from 22-gauge stainless-steel wire insulated except for a .5-mm., sharpened tip. Monopolar electrodes were placed bilaterally in the brain. At autopsy, histological examination of the four brains revealed that in Cats 1 and 2 the electrode tips were placed bilaterally in the anterior hypothalamus; in Cat 3, one electrode tip was located in the left median forebrain bundle in the area of the ventromedial hypothalamus, and the other tip was in the right median forebrain bundle in the area of the dorsomedial hypothalamus. Cat 4 had bilaterally symmetrical electrode placements in the median forebrain bundle in the area of the dorsomedial hypothalamus.

In Cats 1, 2, and 3 there was a 4-mm. distance between electrode tips. There was a 3-mm. distance between electrode tips in Cat 4.

A shuttle box, illustrated in Fig. 1, was used for all conditioning experiments. The apparatus was so constructed that *S* could be introduced into the box from either end. Upon entrance into the box, *S* was placed in a modified dog harness to which the electrode leads were fastened. The leads were then soldered to terminals located on *S*'s head, thus allowing freedom of movement within the box. A sound stimulus could be presented in either compartment of the box through loudspeakers mounted on the doors at each end. Both sound and hypothalamic stimulation could be remotely

controlled from a room adjoining that in which the shuttle box was located. After *S* was placed in the apparatus, it was left alone in the room and the subsequent observations of *S*'s behavior were made through "one-way" viewing screens.

Following recovery from the electrode implantation procedure, the response to hypothalamic stimulation was tested in each *S*. The *S* was placed in the apparatus and hypothalamic stimulation was administered to ascertain thresholds (voltage and current) and to judge the appropriateness of the emotional arousal pattern. It was predetermined, in accordance with our conception of the conditioning process, that only those *S*s would be accepted for the experiment whose response to hypothalamic stimulation included (*a*) emotionality, i.e., hissing, spitting, pilo-erection, pupillary dilation, etc., and (*b*) locomotor behavior, i.e., apparently aimless jumping and running of the kind that crossing the barrier separating the two compartments of the shuttle box was likely to occur.

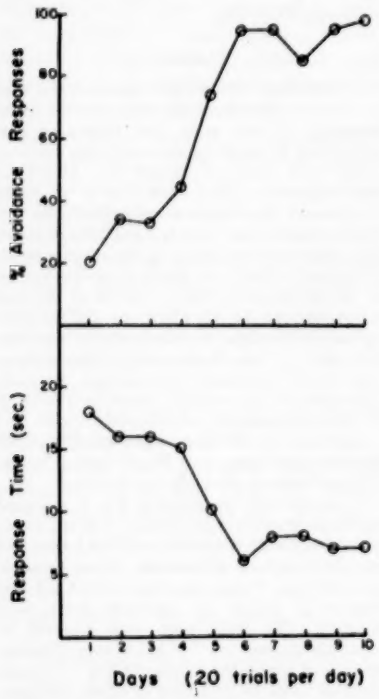

FIG. 2. Mean percentages of avoidance responses and means of the median response times on successive training days ($N = 4$).

The potential and current (all RMS) required for a satisfactory hypothalamic response in Cats 1 through 4 were: volts—6.5, 4.0 2.5, and 2.5, respectively; milliamperes—2.2, 1.2, .73, and 1.1, respectively. The voltage required remained constant for each *S* during the 10 days of experimentation. In all cases a sine wave stimulus of 5,000 cps was used (2).

Each *S* was then given 10 adaptation trials in which a 15-sec. CS was presented at intervals of about 2 min. The CS was a "warbling" tone produced by a Clough-Brengle Audiomatic Audio Generator which was set to sweep through the audio frequency range (0–15,000 cps) five to six times per sec. All *S*s had "adapted" to the experimental situation, including the CS, by the last three trials; they showed no special reaction to the CS and had settled down to one or the other compartment of the box.

Instrumental conditioning was started following the adaptation trials. Twenty trials per day were administered over a period of 10 days, making a total of 200 trials for each *S*. The intertrial intervals were approximately 2 min. each. A delayed conditioning procedure was employed. Each trial was initiated by the onset of the CS. The CS continued for 15 sec., at which time the UCS (hypothalamic stimulation) was automatically administered. Both the tone and the hypothalamic stimulus were continued until *S* crossed the barrier into the opposite compartment of the apparatus. Both tone and shock were simultaneously terminated by *E* as soon as *S* dropped across the barrier. If *S* crossed during the 15-sec. tone period, the tone was terminated by *E* and hypothalamic stimulation omitted. An "escape" response was any hurdle crossing which occurred during direct hypothalamic stimulation (i.e., after the 15-sec. tone alone period). A response was counted as "avoidance" if it occurred during the 15-sec. tone period.

A "shuttle" procedure was employed. Each trial was begun from whichever side of the box *S* occupied at that time. If *S* had crossed to the "left" compartment on one trial and settled down between trials, he was expected to cross to the right compartment on the following trial. What was to be learned was the response of crossing the barrier, rather than a particular "safe" place. The tone always came on in the compartment occupied by *S* at tone onset.

Spontaneous crossing between trials was permitted. This did not occur with sufficient frequency to obscure the conditioning data.

RESULTS

The mean percentage of avoidance responses for each day are presented

in the upper half of Fig. 2. The curve rises slowly from Day 1 through Day 4, and then rises sharply between Day 4 and Day 6, after which it remains fairly stable through Day 10. During the period of rapid acceleration, between Trials 80 and 120, the mean percentage of avoidance response rose from 44% to 95%. On Day 3, three of the four *S*s showed 20% or less avoidance responses. On Day 6, 90% or more avoidance responses were evoked in all four *S*s. On Day 10, three of the four *S*s achieved 100% avoidance responses, crossing the barrier 20 times out of 20 trials without requiring hypothalamic stimulation. The fourth *S* avoided 90% of the time.

The points plotted in the lower half of Fig. 2 represent group means derived from the median response time (in seconds) for each *S* on each experimental day. It is noteworthy that up until Day 4 the curve remains at or above 15 sec. Between Days 4 and 6 the latencies show a sharp drop corresponding to the sharp rise exhibited by the percentage of avoidance.

Qualitatively, the "escape" response appeared to become more effectively directed during the first three days. When hypothalamic reinforcement was necessary, after the third day, the escape response was, in general, highly efficient. In addition, signs of emotionality were noted in response to the CS. The most consistent signs were generalized agitation, crouching and "searching" movements, increased respiratory rate and vocalization.

Discussion

The present results attest to the effectiveness of hypothalamic stimulation as an "unconditioned stimulus" in a typical instrumental avoidance conditioning situation. The instrumental conditioning process appeared similar to that described for conventional avoidance training (**7**). Hypothalamic stimulation produced emotionality and its termination reinforced escape behavior. The CS (the tone) acquired the capacity to evoke emotionality by means of the pairing of tone and hypothalamic stimulation. Finally, termination of the tone and, presumably, its conditioned emotionality, reinforced avoidance behavior.

The central stimulation may be interpreted as a drive-arousing operation resembling the conventional grid-shock technique insofar as the purely motivational (energizing, nonassociative) aspects of the emotionality produced by both are similar. There is, however, an important difference between the two operations. The grid shock produces pain which in turn leads to the arousal of emotional motivation, the termination or reduction of which can serve as an adequate reinforcing condition. Central stimulation enters the sequence of events one step later than the grid-shock technique; emotionality is directly aroused.

The absence of the grid-shock-pain step in the central stimulation sequence can render extremely difficult the selective phase of the instrumental conditioning process. In the conventional grid-pain situation, *S*'s unconditioned withdrawal-reaction tendencies to pain narrows down the response selection problem by providing "built-in" instrumental, pain-terminating (escape) responses. With central stimulation the direction-giving pain cues are absent. Hence it is the responsibility of *E* to choose some discriminable response which regularly occurs during the centrally produced emotional excitement and to terminate the central stimulation when this response is evoked. This was achieved in the present study (*a*) by initially ascertaining that hurdle-crossing locomotion would regularly accompany the emotionality produced by the hypothalamic stimulation, and (*b*) by promptly terminating the central stimulation when hurdle-crossing appeared.

In the experiment by Delgado, Roberts, and Miller (4) the selection process was narrowed down by the pretraining technique mentioned earlier. Masserman's negative results indicated to him that the UCS, in order to be effective, had to be "perceived as noxious" which was apparently the case with conventional grid-shock motivation. It may be correct to conclude with Masserman that the pain-produced emotionality included "meaningful" directional components while the centrally produced hypothalamic emotionality did not. But, it would be incorrect to conclude that the energizing properties of the hypothalamic emotionality could not be so harnessed by appropriate experimental arrangements as to make possible the conditioning of an instrumental avoidance response.

Results recently reported by Olds and Milner (8) suggest that stimulation of certain limbic areas produces approach behavior as defined by the maintenance of high rates of lever pressing in a Skinner Box arrangement. In this situation, central stimulation was automatically delivered to *S* each time the lever-pressing response occurred. The present finding of avoidance conditioning would appear antithetical to such a result. Perhaps the two sets of results can be reconciled. In the lever-pressing situation, it is possible that the reinforcing event is, as in the present study, the decrease in internal stimulation which follows the lever-pressing response. The fact that *S*s do not refrain from approaching the lever in view of the impending stimulation increase which always precedes the final decrease may be due to the absence of direction-giving pain cues connected with centrally produced increments in motivation. A situation of this type conforms with the requirements for "pleasure-seeking" behavior as discussed by Brown (3).

As a final point, it is strongly emphasized that the present results are not interpreted as signifying that the hypothalamus is a "center" for emotion or motivation (6). There are probably many other areas of the brain that could produce motivational emotionality. In addition, it is certain that a complex set of physiological events is simultaneously aroused when the hypothalamus is stimulated and it would be absurd to specify the hypothalamic event as both necessary and sufficient to arouse the resultant motivation or emotionality.

Summary

In order to study the motivational properties of emotional responses produced by direct hypothalamic stimulation, four cats were prepared with permanently implanted hypothalamic electrodes. Each *S* was given 10 days of avoidance training at 20 trials per day. The CS was a warbling tone which sounded for 15 sec. prior to onset of the UCS (hypothalamic stimulation). The UCS was of sufficient intensity to produce strong emotionality ("sham rage") and locomotor behavior. An avoidance response was defined as hurdle-crossing in a shuttle box during the 15-sec. tone period, in which case hypothalamic stimulation was omitted. By Day 6 each *S* exhibited at least 90% avoidance responses. The group mean percentage of avoidance responses rose from 20% on Day 1 to 97.5% on Day 10.

The directional and motivational effects of conventional aversive drive-arousing techniques were compared with the central stimulation technique viewed as a drive-arousing operation. Emphasis was placed on the directional significance of pain cues in the conventional procedures, and the lack of such cues when central stimulation is employed.

References

1. Bard, P. A diencephalic mechanism for the expression of rage with special reference to the sympathetic nervous system. *Amer. J. Physiol.*, 1928, **84**, 490–515.
2. Brown, G. W., & Brown, M. L. Subcortical stimulation at frequencies ranging from 1000 to 15000 cycles per second. *Fed. Proc.*, 1955, **14**, 21.
3. Brown, J. S. Pleasure-seeking behavior and the drive-reduction hypothesis. *Psychol. Rev.*, 1955, **62**, 169–179.
4. Delgado, J. M. R., Roberts, W. W., & Miller, N. E. Learning motivated by electrical stimulation of the brain. *Amer. J. Physiol.*, 1954, **179**, 587–593.
5. Masserman, J. H. Is the hypothalamus a center of emotion? *Psychosom. Med.*, 1941, **3**, 3–25.

6. MEYERS, R. The nervous system and general semantics. IV. The fiction of the thalamus as the neural center of emotions. *ETC. Rev. gen. Semant.*, 1950, **7**, 104–127.

7. MILLER, N. E. Learnable drives and rewards. In S. S. Stevens (Ed.) *Handbook of experimental psychology*. New York: Wiley, 1951.

8. OLDS, J., & MILNER, P. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *J. comp. physiol. Psychol.*, 1954, **47**, 419–427.

9. RANSON, W. W., & MAGOUN, H. W. The hypothalamus. *Ergebn. Physiol.*, 1939, **41**, 56–163.

(Received February 10, 1956)

Journal of Experimental Psychology
Vol. 53, No. 4, 1957

AN ANALYSIS AND REVISION OF HULL'S THEOREM 30[1]

A. C. PEREBOOM[2]

University of California, Los Angeles

In the last of Hull's works, Theorem 30 was introduced for the specific purpose of including the fact that a rapid shift in performance sometimes follows a discrete shift in the magnitude of a reward during a maze-learning process The classical Crespi experiment and the Tolman and Honzik study formed the empirical foundation for his theorem (**2,** pp. 140–148).

The theorem introduces a new habit into Hull's system—a K-habit—which is independent of ${}_{S}H_{R}$. The strength of ${}_{S}H_{R}$ is independent of reward magnitude as long as some reward is present (Postulate VIII[3]), but the K-habit is not. It has two distinct properties: It modifies the size of K—the incentive component of ${}_{S}E_{R}$—by going into operation only when the reward is shifted from one magnitude to another; it accumulates much faster than ${}_{S}H_{R}$, most of the learning occurring within the first two trials after the shift.

Hull's method for computing the new values for K under this theorem can be written

$$K = K_1 + (K_2 - K_1) \times (1 - 10^{-sN_2}), \quad (1)$$

(**2,** p. 143) where K_1 equals the first reward, K_2 the second, N_2 the number of trials under the second reward, and s equals the new rate constant.[4]

The purpose of the present experiment is twofold: Will the theorem predict rapid performance changes in a learning situation requiring the simultaneous discrimination of differential reward magnitudes? Is the K-habit a true habit, a permanent habit like ${}_{S}H_{R}$; that is, is it irreversible and suppressed only by the complete removal of reward?

METHOD

Sixteen male hooded rats 130 to 145 days of age were maintained on a 22-hr. hunger drive throughout the experiment. They had free access to water except during training proper. The order of running of *S*s was randomized from day to day. Following a day's run, a period requiring not less than 1 hr., *S*s were returned to the stockroom, fed for 1 hr., and kept there until the next day's run.

The apparatus was a simple T maze with three nonretrace doors partitioning off the starting section and arms of the T from the stem. A 30-mgm. Friskies pellet was placed on one side of the T and five such pellets on the opposing side. They were in food cups and were not visible from the choice-point area. A noncorrection procedure was used. The first training day followed maze adaptation. It introduced *S*s to their initial training conditions. It consisted of one forced trial to each side. Each of the next nine days consisted of four massed trials: two free, then two forced, with four permissible run patterns—RRLL, RLLR, LRRL, and LLRR.

These 10 days satisfied the learning criterion, a mean of 95% large reward, free trial choices for two consecutive days, Days 9 and 10. Half of the *S*s whose large reward was on the right, and half whose large reward was on the left,

[1] This article is based upon a dissertation submitted in partial fulfillment of the requirements for the Ph.D. degree in Psychology at UCLA. The author wishes to express his indebtedness to Professor John P. Seward whose help and guidance made the research possible.

[2] Now at Texas Technological College, Lubbock, Texas.

[3] This was Postulate IX. It became Postulate VIII in 1952.

[4] Note how this equation contradicts Postulate VIII. ${}_{S}H_{R}$ *must* have been a function of the initial reward magnitude if some representation of that magnitude is to operate on performance when that reward is no longer physically present. See Seward (**6**).

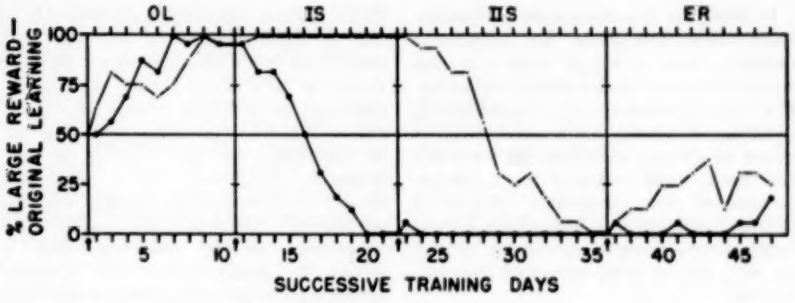

FIG. 1. Daily group percentages of free responses to the side of the T containing the large reward during the original learning (OL). Each vertical division of the graph (↑) indicates a day introducing the subsequent reward conditions. OL: Original learning with unequal rewards. IS: Rewards switched only for Group I (heavy line). IIS: Rewards switched only for Group II (light line). ER: Equal rewards for both groups.

were then randomly assigned to Group I; the remaining eight *S*s comprised Group II. Group I had their rewards switched on Day 11, while Group II found no such change until Day 22. "Switching" consisted of increasing the one-pellet side to five pellets and reducing the five-pellet side to one pellet. As in the case of Day 1, Days 11 and 22 consisted of only one forced trial to each side for both groups. These two days were used to introduce the respective groups to the switch in rewards. All other days consisted of four massed trials: two free, two forced, with the same permissible run patterns as those used during the initial training.

Both groups received a total of 35 training days under unequal rewards. Group I had 10 days under the initial reward conditions and 25 days under the switched conditions. Group II had 21 days under the initial conditions and 14 under the switched conditions. This completed the first phase of the experiment and permitted an evaluation of the generality of the "rapid learning" feature of the theorem.[5]

The second phase followed the first without any break in the continuity of training. At the end of the first 35 days, both groups had reversed their initial preferences; they were now given 12 more training days but under symmetrical rewards. Half of each group now had one pellet on each side of the T and half had five pellets. As in the case of Days 1, 11, and 22, Day 36 consisted of one forced trial to each side while the remaining 11 days consisted of the usual four massed trials.

[5] This first phase parallels the design of an experiment performed by Logan (3) who was investigating differential delays in reinforcement.

RESULTS AND DISCUSSION

Figure 1 shows the development of the initial preference by each group (OL), the reversal of those preferences with reward switching (IS and IIS), and the decay of the reversed preferences with reward equalization (ER).

Reward switching.—If Hull's 1952 system is applicable to our design, then the theorem went into operation, and the K-habits began to develop when the rewards were switched following the development of the initial preference. Since the theorem makes no distinction with respect to the learning itself between an increase and decrease in K (see Equation 1), a rapid learning process should have occurred to both arms of the T, resulting in a rapid reversal in the relative strengths of the respective ${}_{S}E_{R}$'s and thus a rapid reversal in preference.

The Crespi runway experiment and the Tolman and Honzik multiple T experiment that Hull used to support his theorem gave complete shifts of "single" response tendencies in one or two trials. If the separate response tendencies to each side of the simple T in the present study had shifted that fast, then the resulting preferences should have shifted accordingly. They did not, as IS and IIS of Fig. 1 indicate.

If stimulus generalization of choice-point cues had been an important variable, then it would have retarded the development of the initial preference. With that preference once established, however, it would have accelerated a return to chance with rewards switched but would have retarded the complete reversal of that preference. Figure 1 does not support this hypothesis either. Both groups show a symmetry about the 50% line in their curves of response reversal.

How might the difference between the present results and Hull's reference experiments be reconciled? One might argue that the initial stage of performance changes in all these experiments is not a pure measure of reward learning but also includes the retarding action of exploratory behavior (**1**, **5**). This exploration would manifest itself in cul entries, retracings, and starting-and-stopping behavior. The distracting effect of a novel environment would therefore put "inertia" into the original rate of response improvement, but to a diminishing extent as learning proceeds. Rapid performance changes when the incentive shift finally occurred would be a reflection of this greater maze familiarity. Such a phenomenon was less likely to appear in the present design because a choice-point response was used with daily RL experiences equalized throughout training by the use of forced trials.

The difference between Group I and Group II in rate of reversal (IS vs. IIS) appears to reflect a difference in the amount of initial training. Group II had 21 OL days and Group I had only 10. Although not significant ($.10 > P > .05$), a two-tailed test of mean "errors" (free choice responses to the now small, formerly large, reward) in reversing this initial preference suggests this explanation.

Rewards equalized.—Reward magnitudes were equalized in the final stage of the experiment (ER) by replacing one of them with its original magnitude. With reward unchanged on one side of the T, little change in its ${}_{S}E_{R}$ with additional trials should occur. But the opposing ${}_{S}E_{R}$ should also remain unchanged; in order to change it, i.e., to return it to its original strength, it would be necessary for its K-habit to lose strength, to decay. If it does not decay, Hull has, so to speak, "run out of habits." Without any explicit arrangement to the contrary, Hull is forced to predict that the reversed preferences that developed under the K-habits will maintain themselves, although the rewards are now of equal magnitude.[6]

The ER section of Fig. 1 shows that both groups lost their preferences to some degree.[7] Group II with 14 days under switched rewards lost more of it than Group I with 25 days under switched rewards ($P < .01$).

These successive changes in preference suggest that successive amounts of reward are successively learned (**4**). To maintain with Hull that learning is a permanent phenomenon, it will now be necessary to extend his theorem and add additional K-habits in order to handle additional changes in reward magnitude. This view, however, leads to an accumulation of habits and requires a revision of the basic habit construct in order to make each such

[6] It may have occurred to the reader that this argument would also hold if we had switched the rewards for the second time, rather than only shifting one of them back to its original magnitude. However, we avoid the possible criticism that a more complex form of learning based upon successive reversals could be developing, so that we would no longer be dealing with a simple habit like ${}_{S}H_{R}$ or the K-habit.

[7] This was true whether we decreased the large reward (making them equal and small) or increased the small reward (making them equal and large). Because half of each group was under the former condition and the remaining half was under the latter condition, the ER section of Fig. 1 does not show their equivalence, i.e., the lack of a statistically significant difference in the amounts of their preferential loss. When replotted on the basis of this variable, the two curves repeatedly cross each other.

habit distinct from past and future habits acquired in the same S-R situation (e.g., a right turn in the present study). This can be accomplished by incorporating the expectancy principle or Hull's $r_G \rightarrow s_G$ into the habit and allowing reward magnitude to determine the nature and/or intensity of this expectancy. Habit might then be written as Seward has done (6): ${}_{S}H_{r_{1G}}$.

But this revision raises an old problem. If reward magnitude is to determine the distinctiveness of such a habit, what happens to the principle of reinforcement in determining the existence of a habit at all? We shall avoid this issue for now by simply placing K back into the habit as a determiner of its r_{1G} while allowing trials, unspecified with regard to reward, to determine the strength of the learning, i.e., the strength of the connection between the stimulus and that particular r_{1G}.

Considering only K and N, ${}_{S}H_{r_{1G}}$ looks similar to Hull's 1943 version of ${}_{S}H_{R}$: ${}_{S}H_{r_{1G_1}} = K_1(1 - 10^{-iN})$. With subsequent changes in K, however, we must make some assumption about how new habits can come to predominate over former ones acquired in the same situation. Tentatively, we might assume that the s_G of a previous habit is the stimulus for the next one: ${}_{S}H_{r_{1G_1}}H_{r_{1G_2}}$. A quantitative statement of this dual habit resulting from a single change in magnitude is implicit in Hull's Theorem 30 (Equation 1). All that is necessary is to take into account the original learning of reward magnitude, i.e., replace K_1 with $K_1(1 - 10^{-iN_1})$ (see footnote 4).

$$
\begin{aligned}
{}_{S}H_{r_{1G_1}}H_{r_{1G_2}} &= K_1(1 - 10^{-iN_1}) \\
&\quad + [K_2 - K_1(1 - 10^{-iN_1})] \\
&\quad \times (1 - 10^{-sN_2}), \qquad (2)
\end{aligned}
$$

where the subscripts specify the first and second reward and the number of trials under each of them. The superscripts i and s specify the respective rate constants. Simplified, this equation reads

$$
\begin{aligned}
{}_{S}H_{r_{1G_1}}H_{r_{1G_2}} &= K_1(1 - 10^{-iN_1})10^{-sN_2} \\
&\quad + K_2(1 - 10^{-sN_2}). \qquad (3)
\end{aligned}
$$

We are now no longer restricted to one change in the reward as is the case with Theorem 30. For n successive changes the equation reads

$$
\begin{aligned}
{}_{S}H_{r_{1G_1}}H_{r_{1G_2}} \cdots H_{r_{1G_n}} &= K_1(1 - 10^{-iN_1})10^{-(sN_2 + tN_3 + \cdots + zN_n)} \\
&\quad + K_2(1 - 10^{-sN_2})10^{-(tN_3 + \cdots + zN_n)} \\
&\quad + \cdots + K_n(1 - 10^{-zN_n}), \qquad (4)
\end{aligned}
$$

where i, s, t, $\cdots$, z are the separate learning rates for each new K. Although such an equation ignores the possibility of "short circuiting" across these habits with successive changes in quantity, it makes this phenomenon possible since it allows the organism to retain all of his past experiences.

Summary

Hull's Theorem 30 says that a rapid learning process occurs under a new and distinct habit when a discrete shift in reward magnitude occurs during a maze-learning process. An experiment was performed to determine whether a rapid performance change would take place when two simultaneous K-habits were operating and whether the K-habit construct has the permanence feature of Hull's classical ${}_{S}H_{R}$.

Sixteen rats were given two free and two forced trials a day for 35 days in a T maze, with a 30-mgm. pellet on one side and five such pellets on the other. Rewards were switched on Day 11 for eight *S*s (Group I) and Day 22 for Group II. Both groups were then given 12 more days under symmetrical rewards.

No abrupt performance changes occurred with reward switching. An attempt to reconcile this fact with the theorem was then presented. In addition, this experiment suggests that if learning is a permanent phenomenon, then with each successive change in reward magnitude, a new habit must develop which displaces the previous habits.

References

1. Glanzer, M. Stimulus satiation: an explanation of spontaneous alternation and related phenomena. *Psychol. Rev.*, 1953, 60, 257–268.

2. HULL, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.
3. LOGAN, F. A. The role of delay of reinforcement in determining reaction potential. *J. exp. Psychol.*, 1952, **43**, 393–399.
4. MEYER, D. R. The effects of differential rewards on discrimination reversal learning by monkeys. *J. exp. Psychol.*, 1951, **41**, 268–274.
5. MONTGOMERY, K. C. The relation between exploratory behavior and spontaneous alternation in the white rat. *J. comp. physiol. Psychol.*, 1951, **44**, 582–589.
6. SEWARD, J. P. Secondary reinforcement as tertiary motivation: a revision of Hull's revision. *Psychol. Rev.*, 1950, **57**, 362–374.

(Received April 6, 1956)

Journal of Experimental Psychology
Vol. 53, No. 4, 1957

CERTAIN RELATIONSHIPS BETWEEN STIMULUS INTENSITY AND STIMULUS GENERALIZATION [1]

WILLIAM HEYMAN [2]

New York University

The aim of this investigation was to determine the effect of stimulus intensity on indicators of habit and of reaction potential and to relate the results to Hull's revised conception of the role of stimulus intensity and its generalization function.

Stimulus intensity dynamism, V, as expounded in the revised postulates (**8**), has two functions. Habit strength (${}_SH_R$), now determined only by number of reinforcements, is multiplied by a V_1 factor (stimulus intensity dynamism during reinforced trials) to produce a new intervening variable, ${}_S\dot{H}_R$. This intervening variable enters into a multiplicative relation with nonhabit factors in reaction potential (${}_SE_R$), including the second intensity function, V_2 (dynamism during response elicitation). As either V_1 or V_2, the dynamism is conceived as a monotonic, increasing function of the logarithm of intensity measured in physical units.

For generalization along an intensity continuum, Hull postulates that V_1 and V_2 operate as above and that generalized habit strength varies inversely with the absolute difference between the logarithms of the intensities of the conditioning and test stimuli. This alters an earlier formulation (**10**) in which V_2 and the difference between training and test stimuli were considered, but no V_1 effect upon ${}_SH_R$ was assumed. Instead, a ratio, $\dot{V}$ (evocation dynamism divided by conditioning dynamism), was presumed to influence effective reaction potential. Within this context, Hull derived five corollaries. After the revision incorporating V_1, the corollaries were retained except for the last.

Since this paper later compares the advantages of the V_1 or $\dot{V}$ formulation, all corollaries will be paraphrased briefly. Corollary I states that the stimulus generalization gradient extending from a weak stimulus toward the strong extreme of the continuum shows increasing negative slope; whereas by Corollary II, the gradient extending from a strong stimulus toward the weak extreme shows decreasing negative slope. According to Corollary III, when training and test stimuli are equally intense, reaction potential varies directly with the absolute stimulus intensity. By Corollary IV, the stimulus generalization gradient based on a weak training stimulus and relatively strong test stimuli is shallower in over-all slope than the gradient based on a strong training stimulus and relatively weaker test

[1] Based upon a dissertation submitted to the faculty of the Graduate School of Arts and Science of New York University in partial fulfillment of the requirements for the degree of doctor of philosophy. The writer wishes to express his appreciation to the members of his thesis committee: Professors Leland W. Crafts, Ralph W. Gilbert, and Presley D. Stout. Special thanks are due Professor Howard H. Kendler, committee chairman. The writer is grateful to the Psychology Department of Brooklyn College (1951–52) for providing laboratory facilities.

[2] Now at Harpur College, State University of New York.

stimuli. Corollary V predicts that, at the midpoint of the test intensity continuum, reaction potential is equal for responses reinforced at either extreme.

Hull cited reaction time studies with humans and instrumental learning in rats as data which impelled him to conclude that response strength varies with stimulus intensity (**10**); magnitude of GSR conditioned to tones (**6**) and force of pull by rats on a runway (**1**) as evidence for stimulus intensity generalization. He found data from two studies (**1, 6**) consistent with the corollaries, although in one study (**1**), Corollary V was contradicted. However, when the separate function of V_1 was introduced (**8**) and expounded anew (**9, 7**), no clear evidence was adduced for effects of stimulus intensity on habit strength.

Recent experiments, not discussed by Hull, include Grant and Schneider's two studies (**2, 3**), each using a factorial design—one varying light intensity, the other tone. Only the latter yielded a systematic or reliable effect of stimulus intensity on response strength, but, as the authors note, this was possibly caused by the unconditioned relationship of tone intensity to GSR. Neither study indicated a CS intensity effect upon strength of conditioning (such as implied by V_1). In the experiment using light (**2**), an unreliable, rough generalization gradient appeared.

Investigating generalization along the size dimension by varying size of white circles, and thus their reflectance, Grice and Saltz (**4**) found support for Corollaries I, II, and V, and suggestive support for IV. Yet the tables in this article do not show a systematic relation between response strength and absolute stimulus intensity, and, contrary to the V_1 hypothesis, *S*s trained with the larger surface show slightly less resistance to extinction than those trained with the smaller surface.

Since the present experiment was undertaken, Kessen (**11**) has reported the use of variations in light intensity in shock avoidance. He found indicators of response strength to vary monotonically with CS intensity during the acquisition phase, but extinction measures were not systematically related to intensity of either the conditioning or the extinction light. Stimulus generalization gradients were not reported.

The present experiment involved a simple instrumental conditioning situation, designed to minimize opportunities for discrimination learning and to restrict stimulus variations to systematic manipulations of stimulus intensity, in order to assess the effects of stimulus intensity upon the response measure of latency and to ascertain the nature of the stimulus intensity generalization function.

Method

Subjects

The *S*s were 80 experimentally naive, male albino rats of the Wistar strain, between 80 and 85 days old at the beginning of preliminary training.

Apparatus

The apparatus, similar to that described by Wilcoxon, Hays, and Hull (**16**), consisted of a three-compartment box, 8 in. high, constructed of ¼-in. plywood, finished with clear shellac. The start compartment, 8 in. long and 5 in. wide, was separated from the second compartment, 7 in. long and 5 in. wide, by a guillotine door which *E* could raise by a pulley device. At the end of the second compartment, a swinging aluminum door, which *S* could push open and upon which appeared the stimulus papers, led to the third compartment, 15 in. long and 12 in. wide, where, during reinforced runs, a .1-gm. pellet was available in a glass coaster.

By means of two microswitches, a relay assembly, and a Standard electric timer, latencies were automatically timed in hundredths of seconds, beginning when *E* raised the guillotine door and ending when *S* pushed the swinging door.

The five visual stimuli consisted of the brightness values 1.5, 2.5, 4.0, 6.0, and 9.0 of the Munsell Neutral (gray) Series, obtained from the Allcolor Co., New York. The values approximated a scale of equal logarithmic intervals of reflectance, with a mean interval of .4 log units. The Munsell paper, its surface protected by a clear Plexiglas pane, was placed on the swinging door.

Two frosted and shaded 100-w. bulbs, seasoned previously and placed above the apparatus, provided equal illumination in the first and last compartments so that no systematic association between light intensity and *S*'s proximity to food could occur. Light readings with a Weston photometer, Model 603, assured the essential consistency of light conditions throughout the experiment.

Procedure

The experiment was conducted in two replications, each requiring 40 *Ss*. For each replication two stimulus papers, Munsell values 2.5 and 6.0, were used during acquisition (reinforced runs) and five papers, Munsell values 1.5, 2.5, 4.0, 6.0, and 9.0, during extinction (unreinforced runs). Thus, for one group of 20 *Ss*, reinforcement occurred with Munsell 2.5; for the other 20 *Ss*, with Munsell 6.0. From each of these groups, five extinction subgroups of 4 *Ss* each were formed, each subgroup assigned to a different Munsell value. Thus the two acquisition groups yielded a total of 10 extinction groups altogether.

To provide for a large number of trials per day and to keep drive (22-hr. food deprivation) constant among *Ss*, a feeding schedule was maintained throughout. The 40 *Ss* were divided randomly into four blocks, each containing one *S* from each of the 10 extinction groups. Block I was run between 8 and 10 A.M. and fed at 10:30; Block II, run between 11 A.M. and 1 P.M. and fed at 1:30; Block III, run between 2 and 4 P.M. and fed at 4:30; and Block IV run between 5 and 7 P.M. and fed at 7:30. A block about to be run was placed in conveying cages similar to the living cages and removed from the housing room before the preceding block was fed. The standard ration was 10 gm. of Purina Laboratory Chow per *S* per day.

Habituation to laboratory surroundings, feeding, scheduling, and handling was carried out over a 10-day period. Each *S* was handled for about 10 min. daily at times corresponding to the schedule above.

Preliminary training.—This was carried out according to the block schedule from Day 11 through Day 13. On the first trial, *S* was placed in the start compartment and permitted to eat a .1-gm. food pellet. Then *E* raised the guillotine door, and *S* could proceed to the goal compartment while the swinging door was kept fully open. In successive trials, the swinging door was progressively lowered until *S* pushed it open from the fully shut position. For all *Ss*, Munsell 4, the middle paper in the series, was on the swinging door during preliminary training.

Acquisition training.—Again within the block schedule, each *S* received 50 food-reinforced trials at the rate of 10 per day, beginning on Day 14 and continuing through Day 18. (Pilot studies had shown relatively stable asymptotes of latencies after 50 trials.) Practice was distributed; *Ss* spent 10–15 min. in the conveying cage between runs.

From each block of 10 *Ss*, five *Ss*, forming the light-trained group hereafter known as Group L, were run to Munsell 6.0; the remaining five *Ss*, forming the dark-trained group hereafter known as Group D, were run to Munsell 2.5. Latency was recorded in every trial, and *S* received a food pellet in the goal compartment.

Retraining.—Since quasi-massed extinction would necessarily entail variations in the number of days between end of acquisition and beginning of extinction, a retraining procedure, similar to that of Wilcoxon, Hays, and Hull, was employed. Immediately preceding his extinction trials, an *S* was given five quasi-massed reinforced runs with his previous acquisition paper in place.

Extinction.—In each block of 10 *Ss*, two *Ss*, one having been trained with Munsell 2.5 and the other with Munsell 6, were extinguished on each one of the five extinction papers. The extinction procedure omitted both food and food dish. Criterion of extinction was three successive runs, each with a latency greater than 10 sec. If, on a given trial, there was no response within 15 sec., *S* was returned to the conveying cage. If, however, *S* reached the goal compartment, he was left there for 10 sec.

Quasi-massed extinction was used. Within the block schedule, *Ss* were selected by random numbers for single-session, animal-by-animal extinction. Between trials, *S* was returned to the conveying cage while the apparatus was reset. On the average, two *Ss* were extinguished per day. Regardless of achievement of the criterion, all *Ss* were given at least 50 trials, but none more than 100.

To minimize olfactory cues, the box was wiped with a creosote solution at frequent intervals during the experiment.

Results

The results are presented according to their bearing on the following problems: stimulus intensity effects on performance (V_2), stimulus intensity effects on habit (V_1), stimulus generalization, and joint effects of dynamism and generalization. Since latencies typically form skewed distributions, and reliability estimates of differences among means of such distributions tend to be biased by heterogeneity of variances, reciprocals of latencies, i.e., response speeds, were analyzed.

Stimulus Intensity: V_2

To determine effects of V_2 on reaction potential, response speeds for the earliest reinforced runs, acquisition training as a whole, and retraining, as well as speed decrement in the first few extinction trials, were considered.

For each *S*, the median response speed (1/sec.) on the first five trials was obtained. Analysis of variance by acquisition intensities, replications, times of day, and their interactions revealed no factor significant at the 5% level. Thus, at this stage, the V_2 factor was not apparent, and the groups were comparable in other respects.

As an indicator of response strength during the 50 acquisition trials, the median reciprocal latency was obtained for each *S*. In each replication, Group L was slightly faster than Group D; mean median reciprocal latency for Group L was .983, not reliably different ($F < 1$) from that for Group D, .947. Nor were there any other significant sources of variation.

In terms of retraining response speed, a significant difference ($P < .05$) was obtained between Groups L and D, which had mean median reciprocal latencies of 1.242 and 1.015, respectively ($F = 4.22$, with 1 and 74 *df*, the latter based on pooled interactions and residual; Bartlett's test had not impugned homogeneity of variances).

Since, for a given *S*, retraining immediately preceded extinction, retraining speeds were examined in relation to the scheduled extinction stimuli to determine whether systematic trends existed which might erroneously be attributed to stimulus intensity in extinction. No such trends emerged; however, *S*s to be extinguished on Munsell 2.5 were, in both acquisition conditions, somewhat slower than those to be extinguished on other intensities. Nor did classification of *S*s by step-distance of subsequent extinction intensity from training intensity show a hierarchy of response speeds which might later suggest a spurious generalization gradient.

Early extinction runs should provide distinctive evidence regarding V_2; on the whole, response speeds should vary directly with stimulus intensity. In order to reduce variations due to replication and individual differences, median response speed of each *S* on the first five extinction runs was subtracted from its median response speed on the retraining runs. This procedure, effectively like a covariance adjustment, produced a speed decrement score. The V_2 hypothesis implies that decrement, thus computed, should vary inversely with intensity of the eliciting stimulus; i.e., greatest decrement should correspond to extinction Munsell 1.5, smallest to Munsell 9. Table 1 shows no such relationship, and, as Table 2 indicates, the variance estimate due to extinc-

TABLE 1

MEANS AND *SD*'s OF DECREMENT SCORES

(Every entry based on N scores, each determined for individual *S*s by subtracting 1/median latency extinction trials 1–5 from 1/median latency retraining trials. All entries reflect multiplication by 100 of reciprocals of latencies in seconds)

Extinction Munsell	Group D: Munsell 2.5			Group L: Munsell 6			Both Groups	
	N	Mean	*SD*	*N*	Mean	*SD*	Mean	*SD*
1.5	8	22.50	32.54	8	84.12	49.55	53.31	52.02
2.5	8	36.50	41.48	8	69.00	46.04	52.75	46.73
4	8	34.00	29.73	8	86.62	41.53	60.31	44.69
6	8	52.62	41.60	8	51.50	50.08	52.06	46.04
9	8	53.62	41.81	8	78.12	50.34	65.87	47.86
All	40	39.85	39.60	40	73.87	49.30		

tion intensities did not exceed that due to error.

Stimulus Intensity: V_1

Extinction as a whole should provide evidence concerning V_1 since the factor implies greater resistance to extinction for Group L than for Group D. Both median response speed for the first 50 extinction runs (the fewest made by any *S*) and number of trials (n) preceding the criterion of extinction were analyzed for the same sources of variation as in the analysis summarized in Table 2. Transformation to log $(n + 1)$ was used for n because subgroup means were proportional to the respective *SD*s. Neither of these analyses on over-all extinction data led to any significant F ratios, but Group L showed unreliably greater response speed and resistance to extinction than Group D.

A further test of V_1 was possible because the experimental design provided two subgroups extinguished at Munsell 4, the value intermediate between their respective training intensities, 2.5 and 6. These groups did not differ significantly.

As Table 2 indicates, acquisition intensity affected speed decrement significantly. This, however, cannot be interpreted as corroborating a simple V_1 factor. The stronger acquisition stimulus yielded greater speed decrement, contrary to the implication of a simple V_1 effect. The bearing of this result on the experimental hypotheses will be discussed later.

Stimulus Generalization

To avoid interfering effects of generalization of inhibition, generalization of excitation must be assessed

TABLE 2

SUMMARY OF ANALYSIS OF VARIANCE OF DECREMENT SCORES

Source of Variation	*df*	Mean Square†	*F*
Acquisition intensities (A)	1	23,154	11.11**
Extinction intensities (E)	4	582	
Replications (R)	1	94	
Times of day (T)	3	2,160	1.04
A × E	4	2,436	1.18
A × R	1	8,934	4.29*
A × T	3	919	
E × R	4	2,978	1.43
E × T	12	2,042	
R × T	3	1,202	
A × E × R × T (error) pooled with second-order interactions:	43	2,083	

† All mean squares include a constant multiplication factor of 10,000.
* Significant at the 5% level.
** Significant at the 1% level.

in early extinction runs. Measures used were speed decrement and percentage decrement, the latter designed to reduce differences between acquisition groups.

The *S*s were classified according to step-distance of the test from the training stimulus, regardless of acquisition condition and of direction of intensity change. Mean speed decrement, based upon 16 *S*s for 0, 2, and 3 removes, and on 32 *S*s for 1 remove, is plotted against step-distance in Fig. 1, which further shows a linear trend, fitted by the method of least squares. This trend is satisfied by the equation,

$$L = 0.452 + 0.200d$$

where L is the mean decrement (loss) and d the distance between the two stimuli in log units of reflectance (.4 per step interval). The t test of the hypothesis that the trend of means has zero slope yields a value of 8.26, sufficient, with 2 df, to reject this hypothesis at the 1% level. Since this test does not consider intragroup variability, the significant slope merely indicates that specification of decrement for a given group is reliably improved by considering d.

To reduce variability between the two acquisition groups, a percentage decrement was computed for each *S*, as follows: 100 (speed decrement/retraining speed). This measure may be regarded as an approximation to the proportion of reaction potential lost. Figure 2 shows median percentage decrement plotted against step-distance. A positively accelerated trend appears, closely fitted by the exponential relation,

$$\text{Median } L_p = 34.85(1.846)^d,$$

where L_p is percentage decrement and d defined as before.

To determine whether a reliable generalization decrement had been obtained, i.e., whether training at one intensity and test at another produced significantly greater percentage decrement than training and test at the same intensity, a nonparametric test, involving χ^2, was made, since the variates, being ratios, were not normally distributed. Within the acquisition and replication classes, two sets of *S*s were distinguished: base *S*s (trained and tested at the same intensity) and generalization *S*s (tested at an intensity different from training intensity). Whether percentage decrement was greater or not greater than that of the base *S* matched for time of day was determined for each of the 64 generalization *S*s. ("Not greater," rather than "smaller," was used as the class opposing "greater" in order to count the one score equal to the base in favor of the null hypothesis.) By a method described by Snedecor (**14**), chi squares were computed for each replication in each acquisition group. From these, a total χ^2 and a pooled χ^2 were obtained for each acquisition group. The heterogeneity χ^2 (total χ^2 − pooled χ^2) for each acquisition group was small enough to be expected about 3 times in 10 by chance.

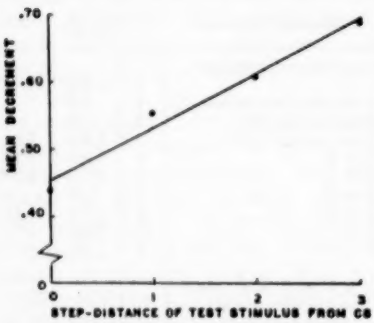

FIG. 1. Mean decrement as a function of distance of test stimulus from CS, in step-distance units of intensity.

Hence, replications could be disregarded in further tests. The acquisition groups were clearly homogeneous; in each, 22 *Ss* had percentage decrements greater, 10 not greater, than the base. For the 44:20 ratio in the experiment as a whole, χ^2, corrected for continuity, was 8.27, which, with 1 *df*, corresponds to $P < .002$, for the one-tailed hypothesis. Thus percentage of response strength lost was greater for generalization than for base *Ss* and did not differ between acquisition conditions.

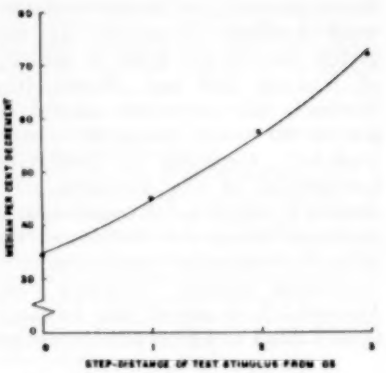

FIG. 2. Median percentage decrement as a function of step-distance of test stimulus from CS.

Joint Effects of Dynamism and Generalization

Simple speed decrement was appreciably greater for Group L (.739) than for Group D (.398); this difference was significant beyond the 1% level. The difference was consistent in direction between replications, although greater in the second than in the first, so that a significant ($P < .05$) A × R interaction was obtained (Table 2). That Group L showed the greater speed decrement seems largely due to its having been tested mostly at intensities weaker than the training stimulus, whereas the opposite shift had prevailed for Group D. This view is supported by lack of direct evidence for a V_1 factor and by differences between the generalization gradients of the two acquisition groups. Neither of these gradients was entirely regular, but that for Group D was shallower. For Group L, the *t* test on correlated means for differences between the origin group (tested on Munsell 6) and the group extinguished at three removes (Munsell 1.5) gave a value of 2.36, with 7 *df*, significant beyond the 2.5% level, according to the one-sided hypothesis tested. For Group D, the *t* between origin (Munsell 2.5) and three removes (Munsell 9) was .93, which could occur 18 times in 100 in repeated random sampling.

Further differences between generalization gradients of the two acquisition groups may be seen when these gradients are reduced to the three-point form which Hull (**10**) devised from available data and related to his corollaries. It should be stressed that such a procedure ignores intragroup variability. For speed decrement, Corollaries I and II imply that the gradient from a weak (here, dark) stimulus toward stronger (lighter) test stimuli should be convex with respect to the X-axis, whereas the gradient from a strong stimulus toward weaker ones should be concave. If we consider Munsell 6, 4, 2.5, 1.5 for the strong-to-weak, and 2.5, 4, 6, 9 for the weak-to-strong gradient, and then, like Hull, obtain a "middle" score for each gradient by averaging the two decrements in the intermediate range, we obtain the gradients in Fig. 3.

These gradients provide technical support for Corollaries I and II, as well as for Corollary IV, which pre-

dicted greater slope for the strong-to-weak gradient. Corollary III, implying that at the point of identity of training and test stimuli (the "origin") the decrement should be greater for weak intensities, is contradicted. Corollary V, concerning intersection of the gradients, also obtains no support. A comparison of gradients within the confines of the intensity continuum common to both acquisition groups (involving only Munsell 2.5, 4, and 6) may be made from Table 1 and gives similar results.

Discussion

The results of the present experiment are interpreted in support of a generalization gradient along the stimulus intensity continuum for simple instrumental responses. The fact that orderliness of the gradient and reliability of the generalization decrement were produced by the use of adjusted scores suggests, consistent with Hovland's findings (6), that variations in intensity make for a shallower gradient than "qualitative" changes in the stimulus.

Hull's theory had asserted that stimulus intensity affects reaction potential directly (V_2 effect) and indirectly through habit strength (V_1). The results offer no indication of an effect on habit; for V_2 only one of several possible indicators yielded significant results. It is recognized that insufficient range of intensities and the presence during all trials of constant incidental stimuli, e.g., tactual ones from the stimulus surface, and relay noises, may have obscured an intensity effect. It is believed, however, in the light of Grant and Schneider's experiments (2, 3) that stimulus intensity probably has a far more limited role than Hull's theory implies; indeed Grant and Schneider's contention—that after conditioning the CS functions as a signal which sets off an existing response mechanism—cannot be dismissed.

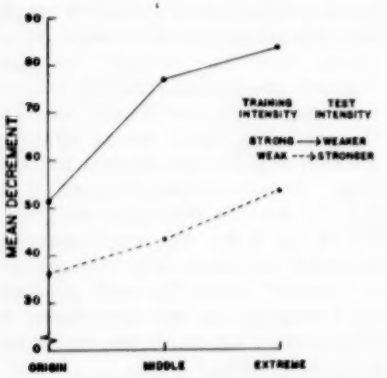

Fig. 3. Three-point generalization gradients. In each, "middle" based on 32 *S*s, each endpoint on 16 *S*s.

Recently, Perkins has presented the interpretation and supporting evidence (12) that the apparent effects of stimulus intensity may be artifacts arising from differential reinforcement of responses to stronger stimuli. Like Skinner (13), he maintains that classical conditioning necessarily entails differential reinforcement, since responses made between trials, in the presence of various constant stimuli differing from the CS, go unreinforced. Inhibition, conditioned to the situation-without-CS, is strongest at the point at which the CS has zero intensity. Net reaction potential is the resultant of gradients of excitation and inhibition, as in Spence's discrimination paradigm (15). Thus, since excitation generalizes symmetrically about the conditioning stimulus and inhibition generalizes symmetrically about zero stimulus intensity, the weakening effect of inhibition will be greatest upon the lower range of intensities.

In instrumental learning situations without differential reinforcement, Perkins argues, apparent intensity effects must be due to special conditions. Thus, in the Grice and Saltz (4) and Brown (1) experiments, since *S*'s approach to the stimulus object was followed by primary reinforcement, a pairing of "perceptually 'more intense'" (12, p. 230) stimulation with proximity of reinforcement may have occurred.

In a basically similar interpretation of Brown's results, the present writer (5)

had suggested an association between the goal gradient and a possible gradient of increasing illumination along the runway, since *S* ran toward the light source. The interpretation of association between "perceptual intensity" of stimuli and gradient of reinforcement would seem not to apply to situations, like that of Grice and Saltz, in which the stimuli are reflecting surfaces under constant illumination; for such surfaces greater proximity would be offset by greater visual angle, so that brightness would remain essentially invariant with *S*'s approach. As noted earlier, Grice and Saltz's data did not show systematic dependence of response strength on intensity. Neither did the present experiment, yet both provided at least suggestive evidence of indirect intensity effects. These will now be considered.

Both Grice and Saltz and the present experiment have provided some support for differences in slope and curvature between gradients of stimulus generalization extending from strong to weak and from weak to strong intensities, respectively. Such differences had been suggested in Hull's Corollaries I, II, and IV. Grice and Saltz further found their results consistent with the implication that the two gradients intersect at the midpoint of the intensity continuum (Corollary V); this may be a function of the particular range of stimulus variation used. Hull seems to have discarded this corollary. The findings common to both studies may be deduced from a modification of Hull's earlier theory (**10**). If the postulate relating reaction potential to absolute stimulus intensity is discarded, but the function relating generalized habit strength to the absolute difference between stimuli, as well as the dynamism ratio (β), as defined earlier, is retained, Corollaries I, II, and IV may be preserved unless further empirical evidence warrants their abandonment.

Summary

This experiment sought to determine the generalization function along the stimulus intensity continuum and to verify whether stimulus intensity is functionally related to indicators of habit strength, of reaction potential, and of the slope and curvature of the stimulus generalization gradient. Thus, the implications of Hull's assumptions concerning stimulus intensity dynamism and stimulus intensity generalization could be tested.

Forty rats were given 50 training trials in the instrumental response of approaching and pushing through a door covered with a light stimulus paper; 40 others were similarly trained with a dark stimulus paper. During extinction, each of these two acquisition groups was divided into five subgroups and tested on the training stimulus or on one of four other stimulus papers. Response speed and speed decrement were the principal measures analyzed.

The major conclusions were:

1. Response strength varies inversely with the difference between conditioning and test stimulus intensities, the form of the function depending upon the measure used.
2. Habit strength does not appear to be affected by the absolute level of the training intensity.
3. Absolute intensity of the eliciting stimulus has little effect upon response strength.
4. The curvature and slope of the generalization gradient may be modified by a factor depending on whether the test stimuli are weaker or stronger than the training stimulus.
5. Most of the experimental findings discussed by Hull in this context, as well as other pertinent data from this and related studies, appear to be susceptible of interpretation in terms of a simple stimulus generalization paradigm and the stimulus intensity dynamism ratio which Hull had suggested earlier.

References

1. Brown, J. S. The generalization of approach responses as a function of stimulus intensity and strength of motivation. *J. comp. Psychol.*, 1942, **33**, 209–226.
2. Grant, D. A., & Schneider, D. E. Intensity of the conditioned stimulus and strength of conditioning: I. The conditioned eyelid response to light. *J. exp. Psychol.*, 1948, **38**, 690–695.
3. Grant, D. A., & Schneider, D. E. Intensity of the conditioned stimulus and strength of conditioning: II. The conditioned galvanic skin response to an auditory stimulus. *J. exp. Psychol.*, 1949, **39**, 35–39.
4. Grice, R. G., & Saltz, E. The generalization of an instrumental response to stimuli

varying in the size dimension. *J. exp. Psychol.*, 1950, **40**, 702–715.

5. Heyman, W. Certain relationships between stimulus intensity and stimulus generalization in the albino rat. Unpublished doctor's dissertation. New York Univer., 1953.
6. Hovland, C. I. The generalization of conditioned responses: II. The sensory generalization of conditioned responses with varying intensities of tone. *J. genet. Psychol.*, 1937, **51**, 279–291.
7. Hull, C. L. *A behavior system: an introduction to behavior theory concerning the individual organism.* New Haven: Yale Univer. Press, 1952.
8. Hull, C. L. Behavior postulates and corollaries, 1949. *Psychol. Rev.*, 1950, **57**, 173–180.
9. Hull, C. L. *Essentials of behavior.* New Haven: Yale Univer. Press, 1951.
10. Hull, C. L. Stimulus intensity dynamism (V) and stimulus generalization. *Psychol. Rev.*, 1949, **56**, 67–76.
11. Kessen, W. Response strength and conditioned stimulus intensity. *J. exp. Psychol.*, 1953, **45**, 82–86.
12. Perkins, C. C. The relation between conditioned stimulus intensity and response strength. *J. exp. Psychol.*, 1953, **46**, 225–231.
13. Skinner, B. F. *The behavior of organisms.* New York: D. Appleton-Century, 1938.
14. Snedecor, G. W. *Statistical methods.* (4th ed.) Ames, Iowa: State College Press, 1946.
15. Spence, K. W. The nature of discrimination learning in animals. *Psychol. Rev.*, 1936, **43**, 427–449.
16. Wilcoxon, H. C., Hays, R., & Hull, C. L. A preliminary determination of the functional relationship of effective reaction potential (sE_R) to the ordinal number of Vincentized extinction reactions (n). *J. exp. Psychol.*, 1950, **40**, 194–199.

(Received April 6, 1956)

Journal of Experimental Psychology
Vol. 53, No. 4, 1957

THE OCULOMOTORIC PATTERN OF CIRCULAR EYE MOVEMENTS DURING INCREASING SPEED OF ROTATION

S. J. GERATHEWOHL, H. STRUGHOLD, AND W. F. TAYLOR

School of Aviation Medicine, USAF

Up to the present time, eye movements in the horizontal plane, such as found in reading, have attracted the particular interest of ophthalmologists, physiologists, and psychologists (**1, 3, 4, 7, 9**). Today, however, the eye is often required to move in other than the cardinal directions because of the technical development and equipment design in special fields of application. In a study of eye movements during radar operation, for example, Gerathewohl recorded circular eye movements when *S* followed the sweep line on the radar plan position indicator (**5**). While eye-movement photographs and some preliminary remarks on the oculomotoric patterns of circular eye movements are given in the study mentioned above, no information is available concerning the circular eye-movement pattern under conditions of varying speed of pursuit. However, it is of interest to know how the eye will behave when it must move along a prescribed circular course at a relatively high speed. The purpose of this study was to record and analyze the basic pattern of guided circular eye movements with regard to increasing velocity. This exploratory study may be useful for the evaluation of oculomotoric time-space patterns and for the design of equipment which requires circular eye movements.

Method

For the experiments described in this paper a Master Ophthalmograph (American Optical Company) was remodeled to take moving pictures as well as still photographs. Details on this type of eye-movement photography have been published elsewhere (**5, 6**). Instead of the reading chart ordinarily used with the ophthalmograph, a silver-plated mirror was placed in the visual field of *S*. In this mirror, the image of a black disc having a diameter of 6 in. was seen at a visual angle of about 10°. A white fixation mark was painted at the periphery of the disc and was illuminated by a 150-w., 120-v. opaque light bulb from a distance of about 40 cm. so that the white fixation mark, which served as the target during the experiments, was easily visible during the rotation of the disc through 360°.

The disc was then rotated at various speeds ranging from 15 to 85 rpm. The *S* was instructed to fixate the white mark and not to lose it while the disc was turning. Ten *S*s (scientists and airmen of the School of Aviation Medicine, USAF) with normal visual acuity for near vision volunteered in the experiments. They were seated before the ophthalmograph and told to adjust their vision so that they could see the whole disc in the mirror. Then the disc was slowly rotated until *S* was familiar with his task, and his eye movements were recorded (*a*) during a constant rotational speed of 15 rpm over a period of 30 sec. At this speed six cycloids or "waves" were obtained for each *S* (see Fig. 1*a*); (*b*) during gradually increasing speed of rotation beginning with a velocity of 20 rpm and ending at approximately 45 rpm. Eleven wave-shaped cycloids were obtained for each *S* during the 30-sec. period; (*c*) during gradually increasing speed of rotation beginning with a velocity of 40 rpm and ending at approximately 85 rpm. This experiment yielded about 24 individual "waves," but not all of them were usable for the analysis because some *S*s were unable to keep their eyes on the fixation mark during the higher rotational speed.

All experiments included clockwise and counterclockwise rotations. Time recordings were obtained on the film by a motor-driven propeller that interrupted the light beam at .5-sec. intervals. The film used in the ophthalmograph during motion pictures moved at a constant speed of 12.5 mm. per sec. so that an experiment lasting 30 sec. yielded a film strip of 37.5 cm. length.

By using the corneal reflection technique, small lights thrown from two lamps of the ophthalmograph are focused on *S*'s eyes. They are reflected by the corneae and then focused on a film. Movement of the eyes causes the light beams to travel and at the same time paint the path of light as bright tracks on the film strip. As the film moves vertically through the camera and *S*'s eyes sweep over the field while fixating the slowly turning target, the perfect pursuit movement is pictured as a cycloid of an almost sinusoidal shape. If the eye stops in its course, the eye fixations are recorded as heavy vertical marks. Saccadic eye movements, on the other hand, appear as deviations from the smooth curve: horizontal movements in the horizontal, diagonal movements in the diagonal, and vertical movements in the vertical plane. Now, with the increasing eye speed superimposed upon the movement of the film, the "waves" lose their smooth and long stretched appearance and approach more and more the shape of a typical pointed cycloid—the arcs representing the upward, and the bends or "troughs" the downward, phase of the rotation cycle of the eye.

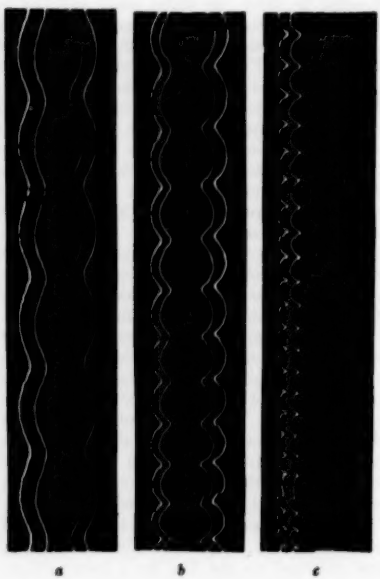

FIG. 1. The oculomotoric pattern of circular eye movements (clockwise, left, and counterclockwise, right) of one *S* at: (*a*) constant rotational speed of 15 rpm, (*b*) speed increasing from 20 to 45 rpm, and (*c*) speed increasing from 45 to 85 rpm.

The records also show that the eye deviates from the prescribed course and that it starts jumping when it cannot pursue continually the fast moving target.

The next problem was to develop a method for the analysis of our records. Number and extent of the saccadic eye movements during this type of pursuit were used as indicators of changes of the oculomotoric pattern. For the quantitative evaluation of the records, the film was magnified by a Recordak film reader having a magnification rate of about 1:15. The graph of either eye was used in counting the number of deviations (saccadic eye movements) and measuring their extension, though usually the clearer line was selected for this purpose. The number of fixations per cycle was first determined, then each saccadic movement was measured in millimeters of deviation from the curve. Since we had found in preliminary experiments that the ratio of up and down movements was about 50:50, the distortion of the size of the jumps caused by the film movement was of no concern. A certain difficulty was encountered, however, in evaluating the records made during increasing rotational speed.

In analyzing these photographs the individual films were matched against standard strips made at constant speeds of 15, 20, 30, 40, 50, 60, 70, and 80 cpm. Since we knew how long a 15-, 20-, . . . 80-cpm cycloid extended on the film, group intervals were established which allowed a sufficiently accurate classification of the individual records. Of course the evaluation would have been easier if each *S* was tested at each of the above-mentioned speeds separately, but this possibility was discarded because better records were obtained when the speed was increased gradually. Some examples are shown in Fig. 1. Since the film speed was known and constant, number and extension of the saccadic eye movements were expressed not only per cycle but also per second of time.

RESULTS

The results given below are based on data of 9 *S*s who were observed at various rotation speeds for both clockwise and counterclockwise rotation directions. One *S* was deleted because he was not observed for all the experimental conditions.

Frequency of eye movements.—Table 1 gives the mean number of saccadic eye movements of the 9 *S*s for each direction and speed of rotation. Table 2 presents an analysis of vari-

TABLE 1

MEAN NUMBER OF SACCADIC EYE MOVEMENTS PER CYCLE AND PER SECOND BY DIRECTION AND SPEED OF ROTATION
($N = 9$)

Cycles per Minute	Number per Cycle				Number per Second			
	Clockwise		Counterclockwise		Clockwise		Counterclockwise	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
15	5.4	3.0	6.5	2.2	1.4	1.0	1.6	0.8
20–24	5.2	1.5	5.6	1.8	2.0	0.8	2.1	0.8
25–29	4.7	1.1	5.0	1.4	2.2	0.6	2.3	0.8
30–34	4.1	1.0	4.4	1.1	2.2	0.7	2.4	0.7
40	3.7	0.9	4.0	0.7	2.5	0.7	2.7	0.6
40–49	3.4	0.8	4.0	1.0	2.6	0.7	3.0	0.8
50–59	3.0	0.7	3.6	1.0	2.8	0.7	3.3	0.9
60–69	2.5	0.7	3.1	1.1	2.7	0.7	3.4	1.2
70–79	2.9	0.6	2.8	0.8	3.6	0.7	3.5	0.9
All speeds	3.8	1.1	4.3	1.2	2.4	0.7	2.7	0.8

TABLE 2

ANALYSES OF VARIANCE OF NUMBER OF SACCADIC EYE MOVEMENTS PER CYCLE AND PER SECOND

Source	*df*	Number per Cycle			Number per Second	
		MS	*F*		*MS*	*F*
Subject (*S*)	8	30.80	$\frac{S}{E}$	34.4**	15.23	34.7**
Rotation speed (R)	8	85.66	$\frac{R}{S \times R}$	17.8**	50.48	39.7**
Direction of rotation (D)	1	47.22	$\frac{D}{S \times D}$	3.41	23.34	3.41
$S \times R$	64	4.803	$\frac{S \times R}{E}$	5.36**	1.270	2.89**
$S \times D$	8	13.84	$\frac{S \times D}{E}$	15.4**	6.839	15.6**
$R \times D$	8	1.746	$\frac{R \times D}{S \times R \times D}$	0.631	1.061	1.31
$S \times R \times D$	64	2.768	$\frac{S \times R \times D}{E}$	3.09**	0.8081	1.84**
Error (E)	450	0.8957			0.4394	

** $P < .01$.

ance for these data. The first thing noticed is that the changes in speed of rotation are apparently associated with changes in the frequency of eye movements. Increasing speed results in fewer such movements per revolution of the target and in more movements per second of time. This result is highly significant statistically. A second result which follows from Table 1 is that a higher number of saccadic movements appears to be associated with a counterclockwise direction of rotation than with clockwise. This effect is not statistically significant, however, since the variation among *S*s is such that chance alone can reasonably account for this result.

The *SD*'s in Table 1 (and later in Table 3) require comment. These are meant to provide estimates of the *SD*'s which would have been observed in the population sampled if each *S* were observed for one cycle (or one second or one movement) only. The figures were obtained by conventional variance component analysis. It was found that the estimates of error variance for each *S* did not vary systematically with rotation speed or direction. The assumption of homogeneous error variance was justified in some of the cases, and all the analyses of variance were done in accordance with this assumption.

In order to simplify the analysis of these data and to obtain orthogonality, a few observations were deleted. Naturally, these deletions were made without regard to the numerical values of the data. Analysis of variance was then performed, using a "mixed" model in a manner analogous to that presented by Bennett and Franklin (**2**, pp. 394–400). The *S*s were considered to be randomly chosen while experimental conditions were fixed. Reference is made to an important paper by Scheffé (**8**) on this subject. Table 2 presents the analysis for number of movements per cycle and per second. The *F* ratios used for hypothesis testing are those indicated by examining the expected mean squares.

It is interesting to note that *S*s differ and that they also show marked differences in response to a change of direction of rotation.

Since the effect of speed of rotation is so pronounced, additional study was made of the relationship between this speed and the number of eye movements. When linear regression lines were fitted to the data, the fit was not very good. A simple sort of curved trend line might describe the findings, but this did not hold true uniformly for all *S*s.

TABLE 3

MEAN EXTENT (IN DEGREES OF VISUAL ANGLE) OF SACCADIC EYE MOVEMENTS PER CYCLE, PER SECOND AND PER MOVEMENT BY DIRECTION AND SPEED OF ROTATION

Cycles per Minute	Extension per Cycle				Extension per Second				Extension per Movement			
	Clockwise		Counter-clockwise		Clockwise		Counter-clockwise		Clockwise		Counter-clockwise	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
15	6.7	5.9	5.4	4.0	1.7	1.9	1.4	1.5	1.1	0.9	0.8	0.9
20–24	7.2	4.4	6.9	4.0	2.7	2.3	2.6	1.7	1.4	1.5	1.2	1.0
25–29	6.1	3.5	9.6	6.3	2.8	1.8	4.4	3.1	1.3	1.0	1.9	1.5
30–34	8.1	4.4	9.3	4.9	4.4	2.8	5.0	2.8	1.9	1.5	2.1	1.3
40	9.4	3.6	11.4	4.7	6.3	2.6	7.6	3.3	2.5	1.3	2.9	1.6
40–49	10.8	4.4	14.8	5.9	8.1	3.5	11.1	4.6	3.2	2.1	3.7	1.9
50–59	12.3	3.6	17.4	6.6	11.3	3.4	15.9	6.1	4.3	2.1	4.8	2.3
60–69	13.2	2.9	18.2	6.9	14.3	3.1	19.7	7.4	5.5	2.7	5.7	3.7
70–79	13.8	3.1	16.7	4.8	17.3	3.7	20.9	5.8	5.0	2.2	6.1	2.5
All speeds	9.7	4.0	12.2	5.3	7.7	2.8	9.8	4.0	2.9	1.7	3.2	1.9

TABLE 4

ANALYSES OF VARIANCE FOR EXTENT OF SACCADIC EYE MOVEMENTS PER CYCLE, PER SECOND, AND PER MOVEMENT

Source	*df*	Extent per Cycle			Extent per Second		Extent per Movement	
		MS	*F*		*MS*	*F*	*MS*	*F*
Subject (*S*)	8	21,920	$\frac{S}{E}$	128**	10.84	102.**	230.6	35.9**
Rotation speed (R)	8	23,885	$\frac{R}{S \times R}$	50.3**	83.38	117.**	1446.	78.3**
Direction of rotation (D)	1	18,822	$\frac{D}{S \times D}$	6.57*	9.304	6.58*	122.9	7.92*
$S \times R$	64	474.45	$\frac{S \times R}{E}$	2.77**	0.7151	6.75**	18.45	2.87**
$S \times D$	8	2863.1	$\frac{S \times D}{E}$	16.7**	1.415	13.4**	15.53	2.42*
$R \times D$	8	2697.6	$\frac{R \times D}{S \times R \times D}$	5.91**	3.083	8.38**	22.09	1.98
$S \times R \times D$	64	456.46	$\frac{S \times R \times D}{E}$	2.66**	0.3680	3.48**	11.18	1.74**
Error (E)	450	171.39			0.1059		6.420†	

* $P < .05$.
** $P < .01$.
† Corrected by dividing by harmonic mean of sample sizes.

Generally speaking, the results of this experiment indicate that the number of saccadic eye movements made while viewing a rotating target are highly individualistic. On the average, there exists a rather regular trend, but individual performances may deviate considerably from this.

Extent of eye movements.—Measurements of the extent of the saccadic eye movements were analyzed in a similar manner. Table 3 illustrates these findings for the extent of eye movements per cycle, per second, and per movement. An analysis of variance study was made of these observations and is reported in Table 4.

When analyzing data for extent of eye movements per movement, it was not reasonable to use the conventional orthogonal analysis of variance. Consequently the analysis was performed on the mean extent per movement obtained by averaging extent of movements for all movements made by a given *S* at a given speed and direction of rotation. A suitable sum of squares for error (SSE) was obtained by dividing the usual within-cell error sums of squares by the harmonic mean of the numbers of movements per *S*-speed-direction cell. This is in accordance with work done by Wilk and Kempthorne (**10**).

Examination of the tables reveals that, as the speed of rotation increases, the extent of deviation of the eye movements from the path of the target increases considerably. This is especially true with regard to the deviations per second. When this is combined with the fact that the number of movements per second increases with increasing speed of rotation, it may be seen that, as the speed increases, the movements of the eye become more frequent and more extensive.

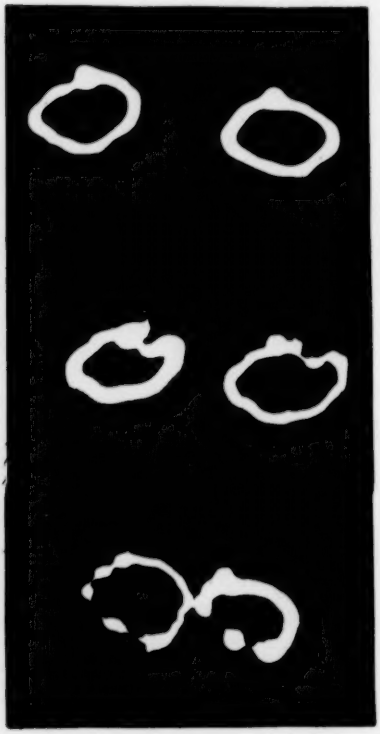

FIG. 2. The oculomotoric pattern of circular eye movements (still photograph) showing an almost perfect (top), a fair (center), and a poor (bottom) pursuit movement at a rotational speed of 15 rpm.

DISCUSSION

Only a few words will be said about the interpretation of these results. Since the study was more of an exploratory nature than a test of a certain hypothesis, its results are not directly applicable to visual perception during motion. On the other hand, it is known that clear vision is lost when the object is moving with an angular speed of more than about 5° per sec. Hence, dynamic visual acuity depends upon linear velocity of the target and is a function of the radius of rotation and—in the present case—of the distance of observation. If the radius is small or the circular path of the target appears at a small visual angle, the circumferential speed of the target is low, and in this case the eye will be able to fixate the object with a minimum of effort. If, on the other hand, the radius is large or is viewed at a close distance, the linear velocity of the target will be relatively high. However, when the eye cannot follow and loses the target, saccadic eye movements occur. Thus the number and the extent of the saccadic movements during visual pursuit indicate *S*'s ability to track the target. It is for this reason that the mean numbers of saccadic movements per cycle and per second were determined and that the extent of the movements made during one cycle and during 1 sec. and the mean extent of the single movement, were computed at the various speeds used.[1]

Looking at Table 1 it is found that the two criteria furnish different information. While the number of saccadic eye movements per cycle decreases from 5.6 to 2.7 clockwise and from 6.8 to 2.9 counterclockwise at the final speed of 70–79 rpm, the number per second increases from 1.4 to 3.4 clockwise and 1.7 to 3.6 counterclockwise at the same velocity. This means that while the number of saccadic movements becomes smaller with the shorter time available for one cycle, the eye must jump more

[1] No attempt was made in this experiment, however, to correlate our two criteria with "dynamic visual acuity."

frequently during each second in order to keep track with the target.

Table 2 shows that the mean extension of saccadic eye movements increases gradually per cycle, per second, and per movement. In this case the average jump covers a distance of about 1° at a rotational speed of 15 rpm and increases to an average of more than 5° at the final velocity of 70–79 rpm. Beyond this speed the data were not clear and reliable for evaluation because only a few *Ss* were able to follow the fast-moving target.

While individual differences were verified in the analysis of variance, the qualitative aspect is strikingly displayed in Fig. 2. These still photographs, taken of three *Ss*, show an almost perfect (top), a fair (center), and a poor (bottom) pursuit movement at a rotational speed of 15 rpm. It became obvious during the experiments that some *Ss* could follow the target up to the highest speed applied, while others were unable to keep the target in sight at speeds higher than 50 rpm. It is not known at the present time to what degree the ability to perform a visual pursuit task can be learned, but there are some indications that *Ss* engaged in ophthalmologic research reached higher speeds than the others. The less experienced *Ss* reported the loss of the target earlier; but again, there is no information on how the increasing jumps impair visual acuity during this type of pursuit.

Finally, the observed differences between clockwise and counterclockwise rotation will be discussed briefly. This effect, being of marginal statistical significance, seems to be consistent enough to permit the conclusion that the counterclockwise rotation of the eyes is more difficult (higher number and larger jumps of the saccadic eye movements) than clockwise rotation. It may be possible that the ease of eye rotation depends upon eyedness or handedness (all *Ss* except one were predominantly right handed and right eyed, and the one left-handed and left-eyed case was inconclusive) and upon the direction in which most of our activities are usually performed.

Summary

The basic pattern of guided circular eye movements was recorded and analyzed with regard to form and speed. Three experiments were made: (*a*) with eye movements at a constant rotational speed of 15 rpm; (*b*) during gradually increasing speeds from 20 to 45 rpm; and (*c*) from 40 to 85 rpm over periods of 30 sec. All experiments included clockwise and counterclockwise rotations. The number and the extent of the saccadic deviations from the prescribed pursuit pattern were used as indices of the accuracy of visual tracking. They were expressed per cycle, per second, and per single movement, respectively.

The data suggests that changes in speed of rotation are associated with changes in the frequency of saccadic eye movements. Increasing speed results in fewer movements per target revolution and in more movements per second of time. A higher number of eye movements occurred during counterclockwise rotation than during clockwise pursuit. An increment of magnitude of saccadic eye movements was found as the speed of rotation increased. By combining the data on number and extension of movements one can say that, as the rotational speed increases, the movements of the eye become more frequent and extensive.

It was concluded that a rotating target can be visually fixated without strain up to a speed of about 30 rpm; that some *Ss* lose pace in the range between 30 and 60 rpm; and that visual pursuit is extremely difficult at speeds higher than 75 rpm. Beyond this limit the oculomotoric pattern disintegrates completely. It is also permissible to presume that moving the eyes in an unusual manner—such as rotation—may cause fatigue and dizziness, particularly with the higher rotational speeds.

References

1. Anderson, I. H. Studies in the eye-movements of good and poor readers. *Psychol. Monogr.*, 1937, **48**, No. 3 (Whole No. 215).
2. Bennett, C. A., & Franklin, N. L. *Statistical analysis in chemistry and engineering.* New York: Wiley, 1954.
3. Clark, B. A camera for simultaneous record of horizontal and vertical movements of both eyes. *Amer. J. Psychol.*, 1934, **46**, 325–326.

4. Dodge, R. The illusion of clear vision during eye movement. *Psychol. Bull.*, 1905, **2**, 193–199.
5. Gerathewohl, S. J. Eye movements during radar operation. *J. aviat. Med.*, 1952, **23**, 597–607.
6. Gerathewohl, S. J., & Strughold, H. Time consumption of eye movements and high speed flying. *J. aviat. Med.*, 1954, **23**, 38–45.
7. Miller, J. W., & Elek, L. Dynamic visual acuity when the required pursuit movement of the eye is in a vertical plane. *U. S. Naval Sch. Aon. Med., Res. Rep.*, Project NM001 075.01.02, 1953.
8. Scheffé, H. A "mixed model" for the analysis of variance. *Ann. Math. Stat*, 1956, **27**, 23–35.
9. Tinker, M. A. The study of eye movements in reading. *Psychol. Bull.*, 1946, **43**, 93–120.
10. Wilk, M. B., & Kempthorne, O. Derived linear models and their use in the analysis of randomized experiments. *WADC Tech. Rep. 55-244*, Vol. II. Wright-Patterson AFB, Ohio: Wright Air Development Center, 1956.

(Received March 2, 1956)

Journal of Experimental Psychology
Vol. 53, No. 4, 1957

DISCRIMINATION LEARNING AND TRANSPOSITION IN CHILDREN AS A FUNCTION OF THE NATURE OF THE REWARD[1]

GLENN TERRELL, JR. · WALLACE A. KENNEDY
AND
University of Colorado *Florida State University*

Several studies have attempted to show the relative effects of different types of rewards in learning problems involving children as *S*s. The problems presented to *S*s in these experiments have been relatively complex. Hurlock (**3**) found that children solve an arithmetic problem more effectively when praised than when reproved, ignored, or treated in a "work as usual" manner. Abel (**1**) found that children trained on a maze problem under delayed-reward conditions made fewer errors than did those who were given praise, a penny, no reward at all, or a combination of these.

The effects of various rewards on children's learning of very simple responses need to be studied. This provides for a more rigorous control of extraneous variables which may be associated either with the experimental (reward) conditions, or the behavior to be learned. Seigel and Foshee (**6**) represent this point of view in their recent study of the law of primary reinforcement in children.

A matter of considerable practical importance concerns the difficulties other *E*s have encountered in selecting appropriate motivational devices when studying children under conditions similar to the present study. Using a small plastic toy as a reward, Kuenne (**4**) found that preschool children required a mean of 206 trials to respond consistently to the larger of two squares. Using a gumdrop as a reward, Alberts and Ehrenfreund (**2**), in what was in many respects a repetition of the Kuenne study, found that preschool children learned the same response in a mean of 47 trials.

All of the previous experiments in this area have employed as a dependent variable some acquisition criterion. It would appear significant to obtain some measure of transposition effect. The purpose of the present experiment was, therefore, to study the relative effects of verbal praise, reproof, candy, and token rewards on the acquisition and transposition of a simple discrimination problem by children. The response learned and transposed was a simple button-pushing response to the larger of two three-dimensional geometric objects.

METHOD

Subjects.—The *S*s were 160 children. There were 80 4- and 5-year-olds, and 80 8- and 9-year-olds. The 4- and 5-year-olds were from five different preschools. The 8- and 9-year-olds were from four elementary schools. The *S*s were randomly selected by schools from all available preschools and elementary schools of Tallahassee, Florida.

Materials.—The stimuli for the problem were three pairs of three-dimensional geometric figures in the shape of cubes, cones, and cylinders (hereafter referred to as the training-stimulus sets). The basal surface area for the three types of geometric figures was the same for each size category. The small member of each stimulus set had a basal area of 4 sq. in., while the large member had a basal area of 8 sq. in. A third cube with a basal area of 16

[1] This paper is based on the doctoral thesis of the junior author under the guidance of the senior author.

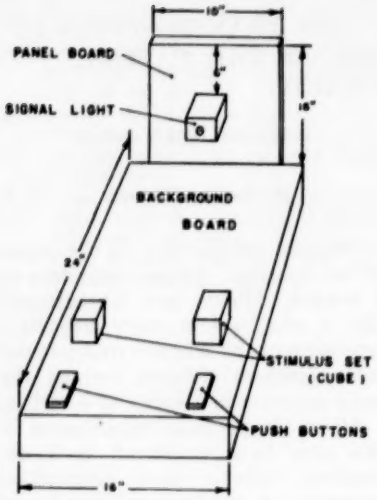

FIG. 1. Schematic drawing of apparatus.

sq. in. was used in a transfer test along with the large cube of the cube-training-stimulus set. This set is hereafter referred to as the test-stimulus set. The stimulus sets were presented to the *S*s unmixed. That is, the large and small cubes were presented together, as were the cones and cylinders. The order of presenting the stimulus sets and the position of the positive stimulus (large) were randomized alike for each *S*.

Additional apparatus consisted of a background board and a panel board (Fig. 1). The background board was a 16 × 24 × ¾ in. piece of plywood with two jacks and two push-button mounts. The stimuli were placed into the jacks on each trial. The push buttons were on a line 5½ in. apart, 5½ in. from the sides, and 1 in. from the front edge of the background board. The jacks were in line with the push-button mounts and were 5½ in. from the front edge. The plywood background board was screwed to the top of a 16 × 24 × 4-in. box which contained the batteries and circuits necessary to operate a signal light. Locked onto the rear edge of the background board was a 10 × 16 × ¾-in. panel board which contained the signal light. This light was mounted in the center of the panel board, 6 in. from the top. The circuits were arranged so that a correct response, pushing the button at the base of the large stimulus, caused the light to go on.

Design.—There were five experimental groups, all of which were rewarded by a light flash after each correct response. The differential reward condition for each group was as follows: The *S*s of Group I, the praise group, were told "That's fine—you are doing well" after each correct response. The *S*s of Group II, the reproof group, were told "No, not that—you are wrong" after each incorrect response. The *S*s of Group III, the candy group, were given a small piece of candy, which they immediately consumed, after each correct response. The *S*s of Group IV, the token group, were allowed to transfer a dried bean from one jar to another after each correct response. The *S*s were told that as soon as all of the beans were transferred to the second jar they could have a small bag of candy. The candy was given to these *S*s immediately upon reaching the acquisition criterion. The *S*s of Group V, the control group, received no reward other than the light flash after each correct response. The experimental design consisted of two levels of chronological age in addition to the five treatment groups. The 4- and 5-yr. children constituted one level, while the 8- and 9-yr. children were the other level. Within each level, *S*s were randomly assigned to the five treatments, making a total of 16 *S*s in each treatment-level combination.

Procedure.—The *S*s were tested individually. Each *S* received the following instructions: "This is a game I want you to try. Choose one of these (*E* points to both the large and small stimulus) and push the button in front of the one you choose. If you are right, that light will go on (*E* points to light on the panel board). If you are wrong, the light will not go on. Now remember, the game is to see how soon you can learn to make the light go on every time." The last sentence was repeated after every tenth trial. Immediately after reaching a criterion of 9 of 10 correct responses on the training-stimulus sets, each *S* was given a 4-trial transposition test on the test-stimulus set. The same differential incentive conditions employed during the acquisition trials were continued during the test trials.

RESULTS

An analysis of variance was made on each of the two criterion measures: number of trials to the criterion, and number of correct responses on the transposition test. The Bartlett test for homogeneity of variance resulted in a nonsignificant chi square on both the training and transposition data. Also, an inspection of both of these distributions revealed no marked non-

TABLE 1

MEANS AND *SD*'S OF TRIALS TO CRITERION IN TRAINING AND CORRECT RESPONSES IN TRANSPOSITION TESTS

(Each treatment group $N = 32$)

Group	Training		Transposition	
	Mean	*SD*	Mean	*SD*
Reward:				
Candy	14.66	4.80	3.91	.29
Praise	26.09	15.04	2.72	1.42
Token	26.53	13.28	3.03	1.44
Reproof	28.16	14.18	2.62	1.59
Control	35.47	17.83	2.78	1.32
Age level:				
4 and 5	28.75	18.15	3.06	1.33
8 and 9	23.61	11.09	2.96	1.44

normality. In view of the results of a study by Norton reported by Lindquist (**5**, pp. 78–86), there seems to be no reason for transforming the raw data.

Table 1 contains both the training and transposition data. As can be seen from this table, the rank order of effectiveness of the five reward groups on the mean number of training trials to the criterion is as follows: candy, praise, token, reproof, and control. The summary of the analysis of variance of the training data appears in Table 2. The interaction is nonsignificant, while the main effects of the reward groups and age levels are significant well beyond the .01 level. The candy group learned the task significantly more quickly than did any of the other groups. No other differences were significant, although the difference between the group that was praised and the one that received only a light flash (control) approached significance in favor of the former ($P = .06$).

TABLE 2

ANALYSIS OF VARIANCE OF NUMBER OF TRIALS TO CRITERION

Source	*df*	Mean Square	*F*
Rewards	4	1784.91	9.65**
Age levels	1	1055.75	5.70**
Interaction	4	335.60	1.91
Within cells	150	184.92	
Total	159		

** $P < .01$.

An inspection of Table 1 reveals the rank order of effectiveness of the five groups on the mean number of correct transposition responses to be as follows: candy, token, control, praise, reproof. The summary of the analysis of variance of the transposition data appears in Table 3. The interaction and age effects are non-

TABLE 3

ANALYSIS OF VARIANCE OF NUMBER OF TRANSPOSITION RESPONSES

Source	*df*	Mean Square	*F*
Rewards	4	8.71	4.98**
Age levels	1	.29	
Interaction	4	2.74	1.57
Within cells	150	1.75	
Total	159		

** $P < .01$.

significant, while again the differences between the treatment-group means reach significance well beyond the .01 level. The candy group transposed significantly more than the praise, reproof, and control groups. The remaining differences between the experimental treatments are nonsignificant, although the candy-token difference approaches significance in favor of the former ($P = .06$).

DISCUSSION

As previously noted, the candy reward was the most effective in both acquisition and transposition, the only exception to this being the candy-token difference on

the transposition test. Since there was no significant interaction present either in training or transfer, this finding applies to ages 8 and 9 as well as to ages 4 and 5. The marked effectiveness of candy as an incentive in this experiment supports a suggestion derived from the Kuenne (4) and the Alberts and Ehrenfreund (2) studies—where it will be remembered that the *S*s of the latter study, rewarded with candy, learned the same response as the *S*s of the former experiment, rewarded with a toy, in one-fourth the number of trials. Apparently *E*s concerned with the problem of selecting the most effective reward for children, an immensely important practical concern, might well consider the results of this experiment. This applies particularly to those who are interested in bringing all *S*s to a training criterion as quickly as possible in order that they may study the effects of certain experimental manipulations on other learning criteria.

An observation of some importance arises from the fact that there were no differences between the older and younger *S*s assigned to the token group either in speed of acquisition or in transposition consistency. For the children trained under this condition the mean number of trials required to reach the criterion were 26.06 and 27.00 for the younger and older *S*s, respectively. The mean number of transposition responses for the younger and older age levels were 3.00 and 3.06, respectively. Though the writer knows of no research on this matter, supervisors and teachers of children frequently report on the ineffectiveness of delayed rewards with very young children. Apparently 4- and 5-yr. children can learn as effectively as 8- and 9-yr. children under delayed-reward conditions, provided they are clearly able to observe progress toward the ultimate goal. It is obvious that this hypothesis needs to be tested in experiments concerned with other types of learning problems and performed with *S*s from the same and different age levels before much confidence may be placed in its accuracy.

Summary

Subjects (80 4- and 5-yr. children and 80 8- and 9-yr. children) were randomly assigned to five reward conditions: praise, reproof, candy, token, and control (light flash). Training to the larger of two three-dimensional stimuli was continued until *S*s had reached a criterion of 9 of the last 10 responses correct. A 4-trial transposition test was then given with stimuli one step removed from the training pair. The candy-reward group learned significantly more quickly than did any of the other groups; while on the transposition test, the candy group was significantly superior to all groups except the token-reward group. The older children learned significantly more quickly than did the younger, but there was no age-level difference on the transposition test.

References

1. Abel, L. B. The effects of shifts in motivation upon the learning of a sensorimotor task. *Arch. Psychol., N. Y.*, 1936, No. 205.
2. Alberts, E., & Ehrenfruend, D. Transposition in children as a function of age. *J. exp. Psychol.*, 1951, **41**, 30–38.
3. Hurlock, E. B. The value of praise and reproof as incentives for children. *Arch. Psychol., N. Y.*, 1924, No. 71.
4. Kuenne, M. R. Experimental investigation of transposition behavior in young children. *J. exp. Psychol.*, 1946, **36**, 471–490
5. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Chicago: Houghton-Mifflin, 1953.
6. Siegel, P. S., & Foshee, J. G. The law of primary reinforcement in children. *J. exp. Psychol.*, 1953, **45**, 12–14.

(Received March 5, 1956)

Journal of Experimental Psychology
Vol. 53, No. 4, 1957

EFFECT OF CONTROL LAG ON PERFORMANCE IN A TRACKING TASK [1]

JACK E. CONKLIN

Minneapolis-Honeywell Regulator Company Aeronautical Division

In a typical tracking experiment, *S* is seated in front of a cathode ray tube (CRT) on which is displayed a target and cursor. As the target continually changes position on the CRT, *S* is instructed to maintain alignment between the target and cursor by appropriate control movements. Skilled performance in this task requires anticipatory behavior on the part of *S*. Predictions about the future position of the target permit *S* to compensate for his own functional intermittency (e.g., reaction time) and produce a response output which is a continuous function of the problem.

The transmission characteristics of the external control system have an influence on tracking accuracy. Hick and Bates (**8**) distinguish among three types of control systems, i.e., position, rate, and acceleration. When cursor position is directly proportional to control displacement, the system is termed a position control. When cursor velocity or acceleration is determined by control position, it is known as a rate or acceleration control system, respectively. With a position control, a step function input will cause the cursor to move instantaneously to a new resting position. With rate or acceleration controls, however, the cursor will not come to rest until the control is returned to its neutral position. Controls may also be designed so that both position and rate are a function of control displacement, viz., rate aided control systems.

The transmission characteristics of a control may also be altered by the presence of lag in the control circuit. Two common varieties of lag with a position control are transmission and exponential lag. Following a control movement, a fixed time interval can expire before information relative to this movement is visually displayed. This is termed transmission lag. Exponential lag, however, produces an exponential output of control displacement. The consequence of a control movement is immediately perceived, in part, but attains completion as an exponential function of time. Conventionally, exponential lag in seconds is defined as the time interval between the control input and 63% of the maximum output.

Previous research on the effects of lag appears to warrant the following tentative conclusions: (*a*) tracking performance is impaired with the presence of relatively small transmission-type delays in the control system, even lags below the threshold of perception; (*b*) exponential lag becomes increasingly detrimental as target course frequency increases; and (*c*) as exponential lag is increased from zero to 2.7 sec., tracking effi-

[1] This investigation was conducted as part of a Ph.D. dissertation submitted to the Graduate School of The Ohio State University. The author wishes to express his appreciation to Dr. Paul M. Fitts for his valuable suggestions and helpful criticisms. The research was supported in part by the United States Air Force under contract No. AF 33(038)-10528 with The Ohio State University Research Foundation, monitored by the Air Force Personnel and Training Research Center. Permission is granted for reproduction, publication, use and disposal in whole or in part by or for the United States Government.

ciency decreases rapidly at first and then more slowly as the delay is increased. These conclusions are all based on tracking performance with a compensatory display.

Transmission-type control lags up to .32 sec. were investigated by Warrick (12). On a compensatory display, the target was set into motion by a mixture of two periodic waveforms (6 and 30 cpm). Time-on-target scores were found to decrease progressively as the delay interval was increased. Attenuation of tracking performance was found to occur with transmission lags as short as .06 sec. Investigation disclosed that the threshold for the perception of control lag was also at this value (1). Lag diminished tracking efficiency even when *S* was not aware of its presence.

In another experiment, Warrick (11) studied compensatory tracking with exponential lag. Three values of control sensitivity (15, 45, and 65 degrees of knob rotation), three target frequencies ($\frac{1}{3}$, $\frac{1}{4}$, and $\frac{1}{5}$ cps), and three lag time constants (.005, .105, and .255 sec.) were employed. Time-on-target scores declined with an increase in target frequency and asymptotic lag. A change in control sensitivity had little effect on tracking efficiency. A significant interaction between frequency and asymptotic lag suggested that as target frequency increased, lag became increasingly detrimental.

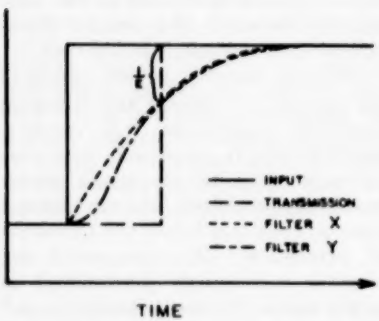

FIG. 1. Illustration of transmission, asymptotic, and intermediate lag function.

In a recent study by Levine (9) exponential lag was varied between .015 sec. and 2.70 sec. with a compensatory display. Target motion was derived by a mixture of three frequencies with equal amplitudes (3, 5, and 11 cpm). Levine observed a linear relation between tracking efficiency and asymptotic lag for values of .15 sec. to 2.70 sec. From 0 sec. to .15 sec., however, the function assumed an exponential shape. He also found that tracking performance was partially successful even with the longest delay, i.e., 2.70 sec. At this value of lag the mean time-on-target measure for 11 *S*s was 13.6 sec. With the control maintained in a neutral position, the time-on-target value of the signal was only 4.2 sec.

The present experiment was designed to determine that value of control lag where skilled performance breaks down. Tracking performance was observed with a pursuit and with a compensatory display, and with four different target courses. An additional experimental variable was the nature of the control-lag function. Lag was introduced by an electrical filter network. Keeping the time to reach 63% of maximum output constant, the function was exponential when one filter network was employed (Filter X) but had a different shape when two networks were cascaded (Filter Y). The function derived with a two-filter network was intermediate between asymptotic and transmission-type lags (see Fig. 1). A response to a step-function input was an initial positive acceleration which changed to negative as maximum output was approached.

METHOD

Apparatus.—The equipment employed in this experiment was the OSU Electronic Pursuit Apparatus. A detailed description of an earlier model was given elsewhere (10). The principle elements consisted of (*a*) target course generators, (*b*) display, (*c*) control, (*d*) timing circuits, and (*e*) recording units.

Two target course generators were used for this study. One furnished three simple harmonic signals, viz., π, $\frac{1}{2}\pi$, and $\frac{1}{4}\pi$ radians per sec. This generator permitted these frequencies to be displayed individually or in various mixtures. Amplitudes were separately controlled by rheostats. A second generator produced a random signal which had peak power at π radians per sec., and a sharp drop to near zero power at frequencies above this value. A complete description of the random course generator has been given in a separate report (4).

The display was presented on a 5-in. CRT located approximately 28 in. from *S*'s faceplane and perpendicular to the plane-of-regard. The target was a thin vertical line $\frac{1}{16}$ in. wide and $\frac{1}{2}$ in. long. The cursor was a second line with the same dimensions and was displayed beneath the target. Room illumination was maintained at a low level in order to increase the contrast between the display and immediate surround. All displayed motion was horizontal, i.e., one dimensional.

The apparatus permitted either pursuit or compensory information to be displayed. The horizontal movements of the target and cursor with the pursuit display were independent. The former motion was activated by the target course generator, the latter by *S*'s control movements. Compensatory information was displayed differently. The upper vertical line (target) remained stationary in the center of the CRT. Movements of the lower line were the result of both the target course input and *S*s control movements. With both displays the tracking task was to maintain a continuous alignment between target and cursor.

Control movements consisted of an angular rotation of the arm about the elbow. The *S* placed his arm on a horizontal lever with his elbow at the pivot point. His hand was then in position to grasp a knob at the other end of the lever. The distance of the knob from the elbow was adjusted to the arm length of each *S*.

Timing circuits controlled the 90-sec. tracking period. A warm-up period of 5 sec. preceded each trial before the recording units were activated. This gave *S* time to acquire the target before measurements of his performance were taken.

Tracking error is defined as the difference between the problem voltage and the voltage determined by the position of the arm control at any point in time. This is perceived by *S* as a displacement between target and cursor. By means of electronic squarers and integrators, tracking error voltages are squared and accumulated for each trial. The total mean squared error voltage then appears on a voltmeter at the end of the tracking period.

Traced reproductions of the control lag values used in this experiment are given in Fig. 2. The time constants employed range between 0 sec.

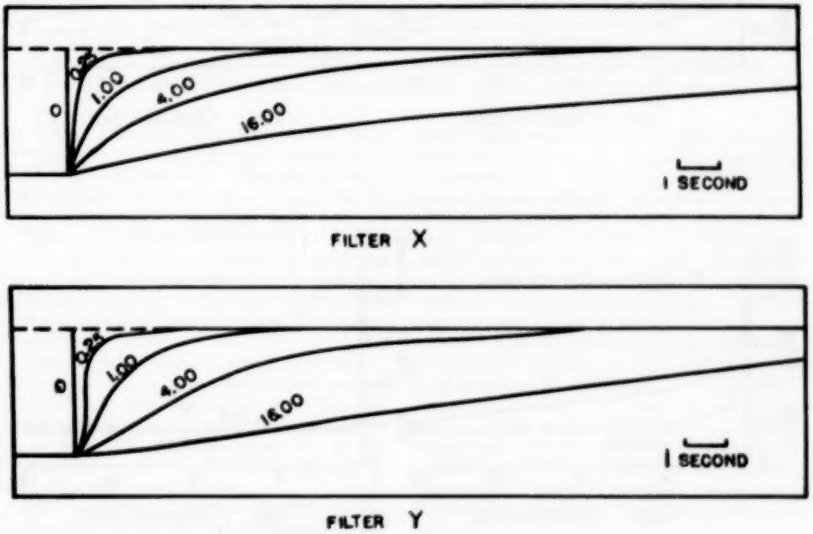

FIG. 2. Traced reproduction of control lag functions from 0 to 16 sec.

and 16 sec. To obtain the desired control lags precisely, it is necessary to employ a trial-and-error procedure substituting resistor values approximately equal to the computed values and calibrating the function with a sensitive graphic recorder.

Subjects.—Data were collected on three skilled *S*s each of whom had several months of previous tracking experience under a variety of conditions.

Experimental variables.—Three repetitive target courses differing in intricacy and one band-limited random course were employed. The repetitive course consisted of (*a*) a simple sine wave (π rad./sec.), (*b*) a mixture of two sinusoidal components ($\pi + \frac{2}{3}\pi$ rad./sec.), and (*c*) three sine waves in mixture ($\pi + \frac{2}{3}\pi + \frac{1}{6}\pi$ rad./sec.). Frequencies were combined with equal amplitudes, but the over-all voltage of each of the signals was adjusted so that peak values were equal. The root mean squared (*rms*) values of the repetitive courses were 44.3, 34.7, and 27.2, respectively. The random signal had an *rms* value of 21.4.

Five conditions of control lag were employed, i.e., 0, .25, 1.0, 4.0, and 16.0 sec. (see Fig. 2). Lag is defined here as the time taken for the output of the filter to reach a value of 63% of its final value following a sudden or step-function input. These values were studied with Filters X and Y. Data were collected with the pursuit and compensatory display.

Procedure.—For a given experimental session, *S* tracked twenty 90-sec. trials with a 30-sec. rest period between trials. The display and control scale factors were held constant at 2 in. and 5°, respectively. Thus the total range of movement of the target covered 2 in., and *S* had to move the arm control 5° as measured at the elbow in order to move the cursor 2 in. on

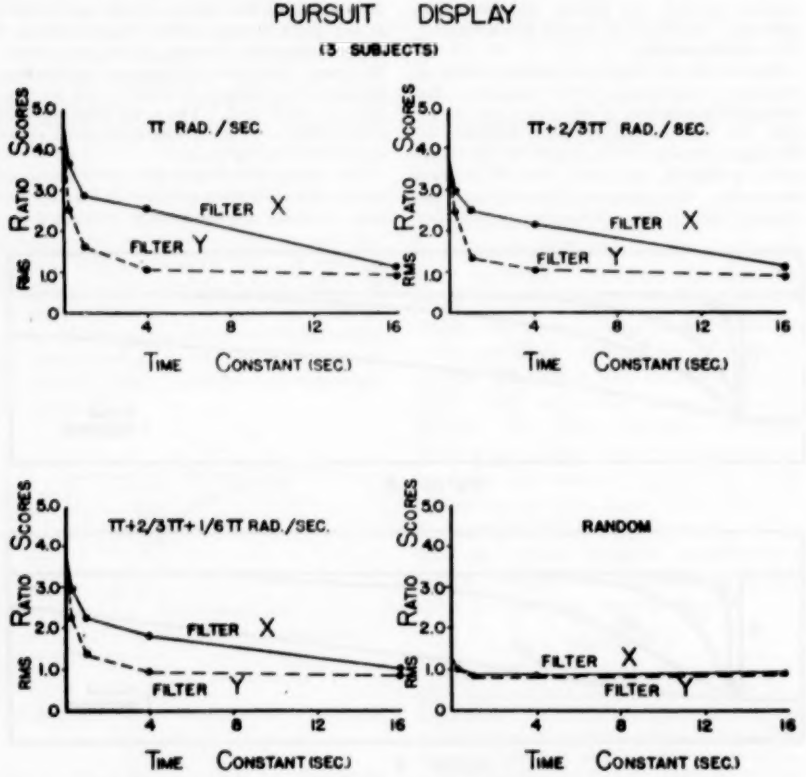

FIG. 3. Tracking efficiency as a function of control lag with the pursuit display.

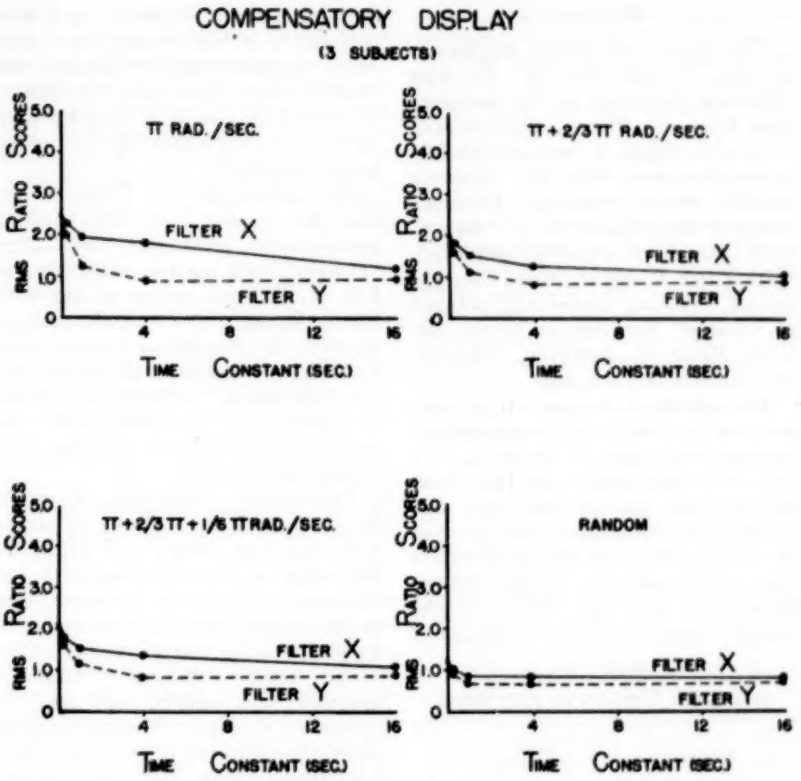

FIG. 4. Tracking efficiency as a function of control lag with the compensatory display.

the CRT. The five conditions of lag were repeated four times during each session in an ABCDEEDCBA sequence. The type of target course and the filter network (Filter X or Y) were held constant for any given day.

Data were collected with the pursuit display for the first eight days, followed by eight sessions of compensatory tracking. The order in which the four target courses and the two filter networks were given was determined by a table of random numbers.

Two practice sessions preceded the part of the investigation in which the pursuit display was used, and two sessions preceded the collection of data with the compensatory display. All *S*s were given equal experience on (*a*) the different lag conditions, (*b*) Filters X and Y, and (*c*) the four target courses.

The tracking criterion.—Inasmuch as the display scale factor was held constant, target course inputs were regulated so that the maximum amplitude for any particular signal did not exceed 2 in. This introduced differences among the *rms* values of the signals. In order to compare tracking performance under different target conditions, therefore, a criterion independent of the *rms* value of the signal was needed. The criterion employed was the ratio of the *rms* error of the signal to that of *S*. This transformation has the advantage of (*a*) equating different target course inputs so that they could be compared directly, and (*b*) determining the level of performance at which *S* is tracking with respect to "chance," i.e., with respect to the score he would obtain if he merely centered the control and held it stationary throughout the tracking period. A ratio score of 1.0, therefore, indicated a "chance" tracking performance. Higher scores revealed *S*'s success in reducing the error.

Results

The results are shown graphically in Fig. 3 and Fig. 4. Tracking efficiency decreased as lag increased from 0 sec. to 16 sec., and performance with Filter Y was consistently poorer than with Filter X. With the random signal, tracking efficiency dropped immediately to a "chance" level with the shortest lag employed. With Filter X and the three coherent signals, tracking proficiency reduced to "chance" behavior at 16 sec. lag. With Filter Y, however, "chance" was reached at 4 sec. lag.

The results of an analysis of variance for the pursuit and compensatory displays are given in Table 1. To test the significance of the main effects and the interactions, the fourth-order interaction was used as the error term. All the main effects were significant, i.e., target course, filters, lag, and *Ss*. Three of the double interactions were significant with the pursuit display, viz., Filters × Lag, *Ss* × Filters, and *Ss* × Lag. With the compensatory display, two double interactions were significant: Filter × Lag, and *Ss* × Target course. None of the triple interactions proved to be significant.

TABLE 1

Analysis of Variance of *RMS* Error

Source	*df*	Mean Square	*F*
Pursuit Display			
Target course	2	2.8829	13.79**
Filters (X and Y)	1	46.6578	223.14**
Ss	2	14.7330	70.46**
Lag	3	40.3795	193.11**
Filters × *Ss*	2	10.0031	23.92**
Filters × Lag	3	10.2381	16.32**
Ss × Lag	6	10.56[illegible]7	8.42**
Error	12	0.2091	
Compensatory Display			
Target course	2	1.9770	30.28**
Filters (X and Y)	1	10.8152	165.62**
Ss	2	2.8975	44.37**
Lag	3	8.9664	137.31**
Filters × Lag	3	0.6866	10.51*
Target × *Ss*	4	0.4304	6.57*
Error	12	0.0653	

* $P = .01$.
** $P = .001$.

The interaction of Filters × Lag was due to the fact that tracking performance reached the same "chance" level (an *rms* ratio score of 1.0) at different values of lag with Filters X and Y. The interaction of *Ss* with the remaining experimental variables presumably reflects individual differences. Differences among *Ss* diminished, however, with the more difficult tracking conditions, i.e., 16-sec. lag, Filter Y.

Performance instability.—Instability in *S* performance was apparent with the long delays and Filter Y. The effect is most striking with the random target course. In attempting to reduce perceived error, *Ss* achieved a performance score considerably worse than "chance." Had *S* centered the control and held it stationary during the difficult tracking tasks, the mean *rms* ratio scores would have been at least 1.0. As it turned out, some of the ratio scores recorded were below .60.

Discussion

The results support the hypothesis that skilled behavior in a perceptual motor task depends to a large extent upon *S*'s ability to anticipate and predict system performance. Furthermore, it is evident that predictions enable *S* to overcome or compensate for his basic intermittency and thus behave in a system as a continuous error-correction device. In a tracking task the basis for prediction arises from two sources: (*a*) long-term predictions founded on what *S* had learned about the target pattern, and (*b*) short-term predictions mediated by the perception of target rates and

direction. Error information per se is of little use as an indicator of future target position. These data suggest that interference with short-term predictions can be accomplished by the manipulation of task variables and that this causes a decrement in tracking proficiency. The differential effect of (*a*) Filters X and Y, (*b*) coherent and incoherent target courses, and (*c*) the pursuit and compensatory display on tracking performance, all lend credence to this view.

The most striking result is the detrimental effect of Filter Y on performance as contrasted to Filter X. Performance decrement is not due to *S*s inability to learn the pattern of target movement, such as the sine wave, but rather to the lack of immediate verification of his short-term predictions. With Filter Y, the *rms* ratio scores drop off more rapidly, and tracking reaches a "chance" level of proficiency at much shorter values of lag. This result holds for the compensatory and pursuit display, and all target courses employed except the random signal. With straight asymptotic lag (Filter X) *S* is given some instantaneous information about control movements—enough initial information to verify, in part, his predictions. With Filter Y, the initial output is zero and the output does not reach a discriminable level for a fraction of a second. Lacking immediate information, predictions become uncertain and tracking performance falls off at a rapid rate as time constants are increased. Predictions, then, appear to be severely limited when *S* cannot perceive, even in part, an immediate effect of his control movement. Birmingham, Kahn, and Taylor (**2**) would discuss these results by their concept of display "quickening." A "quickened" display is one which presents to *S* immediate and complete information about his corrective responses. Displayed information with Filter Y is presumably less "quickened" than with Filter X.

That anticipatory behavior on the part of *S* is essential for good tracking performance is also indicated by the results with the random signal. Since the random pattern is nonrepetitive, tracking performance is primarily dependent upon short-term predictions revealed by moment-to-moment target rates. This makes it possible for a skilled *S* to achieve a slightly better than "chance" performance when there is no delay in the control circuit. As soon as a short lag (i.e., .25 sec.) is introduced, however, predictive ability is destroyed and tracking performance breaks down. These results further emphasize that when predictions become impossible, *S* can no longer perform as a stable element in the system. Any effort to track on his part is likely to lead to an actual increase in error beyond the value it would reach if he did nothing.

Finally, the assumption that anticipatory or predictive behavior is essential for adequate tracking performance is further supported by the consistent differences found between pursuit and compensatory tracking. Under comparable conditions of lag and intricacy, performance with the compensatory display is always inferior. Difficulty arises in using the compensatory display as a basis for predictions, since the motion of the target is confounded with control movements. It is very difficult for *S* to differentiate these two effects. In addition, *S* receives no information about target rates or direction of motion when cursor and target are aligned. He must depend primarily upon displacement cues which give little, if any, predictive information. Perfect alignment of target and cursor with the pursuit display, however, does not impair predictions. Target rates and directions are clearly perceptible, and they present informational cues essential for skilled performance.

That *S* is capable of behaving as a continuous error-correction servo is in contrast to the prevailing view (**3, 5, 6, 7**). Birmingham and Taylor (**3**, p. 7), for example, emphasize the intermittency in the human response and neglect the important role of intellectual factors in skilled performance. This approach

delegates to *S* the simplest function, e.g., that of a linear amplifier, and loses sight of human versatility of performance. That man-machine systems could be designed for efficient operation by curtailing the role of the human element is seriously questioned by the writer. More research is needed in this area.

An empirical study of control lag has a significant practical application for the design of control systems. In radar gunnery, for example, engineers are frequently faced with the problem of noise (target "jitter and wander") on the radar screen. Electrical filters may be employed to remove noise components from the system. It is important, therefore, to determine how much the human operator can be taxed in his predictive behavior before the system becomes unstable. The evidence indicates that asymptotic lags of the order of 16 sec. are the upper limit with intricate target courses containing frequencies as high as π radians per second. With simplified target motion, however, the human operator may overcome the effects of greater lags and continue to perform as a stable element in the system. Electrical and mechanical systems are commonly constructed with two or more lag effects in cascade. Under these circumstances, relatively short time-constants will have a serious effect on operator stability as shown by this experiment. Further studies are needed in this area to evaluate the effect of cascading more than two RC networks on tracking performance.

Summary

The experiment is concerned with the effect of delay between the control and display on tracking activity. Lag was introduced by two different filter networks—a single RC network (asymptotic lag) and two RC networks in cascade. With a pursuit and compensatory display, four different target courses were studied with delays of 0, .25, 1.0, 4.0, and 16.0 sec. The results supported the assumption that predictive behavior on the part of *S* is essential for skilled performance. Experimental task variables that impair short-term predictions about future target position seriously limit tracking efficiency.

References

1. Biel, W. C., & Warrick, M. J. Studies in the perception of time delay. Unpublished, Aero Medical Laboratory, Wright Field AFB, Dayton, Ohio, 1949.
2. Birmingham, H. P., Kahn, A., & Taylor, F. V. A demonstration of quickening in multiple-coordinate control tasks. Naval Research Laboratory, Washington, D. C., *NRL Rep.* 4380, 1954.
3. Birmingham, H. P., & Taylor, F. V. A human engineering approach to the design of man-operated continuous control systems. Naval Research Laboratory, Washington, D. C., *NRL Rep.* 4333, 1954.
4. Clark, J. R., Fontaine, A. B., & Warren, C. E. The generation of a continuous random signal for use in human tracking studies. Human Resources Research Center, Lackland AFB, *Res. Bull.* 53–40, 1953.
5. Craik, K. J. W. Theory of the human operator in control systems. I. The operator as an engineering system. *Brit. J. Psychol.*, 1947, **38**, 56–61.
6. Craik, K. J. W. Theory of the human operator in control systems. II. Man as an element in control systems. *Brit. J. Psychol.*, 1948, **38**, 142–148.
7. Hick, W. E. The discontinuous function of the human operator in pursuit tasks. *Quart. J. exp. Psychol.*, 1948, **1**, 36–51.
8. Hick, W. E., & Bates, J. A. V. The human operator of control mechanisms. Great Britian, Ministry of Supply, *Mono.* No. 17-204, 1950.
9. Levine, M. Tracking performance as a function of exponential delay between the control and display. USAF Wright Air Development Center, *Tech. Rep.* 53-236, 1953.
10. Warren, C. E., Clark, J. R., & Fontaine, A. B. A two-dimensional electronic pursuit apparatus. Human Resources Research Center, Lackland AFB, *Res. Bull.* 52-26, 1952.
11. Warrick, M. J. Effect of controller lag, gear ratio, frequency of pointer oscillation and their interaction on compensatory tracking. Mimeographed, Aero Medical Laboratory, Wright Field AFB, Dayton, Ohio, 1950.
12. Warrick, M. J. The effect of transmission type control lags on tracking accuracy. USAF Air Materiel Command, *Tech. Rep.* 5916, 1949.

(Received March 12, 1956)

Journal of Experimental Psychology
Vol. 53, No. 4, 1957

NONREINFORCEMENTS VERSUS REINFORCEMENTS AS VARIABLES IN THE PARTIAL REINFORCEMENT EFFECT

HOWARD H. KENDLER, STANLEY S. PLISKOFF,
MICHAEL R. D'AMATO, AND SANFORD KATZ

New York University

The reasons why a partial reinforcement schedule usually produces greater resistance to extinction than a continuous reinforcement schedule has been the source of much theoretical speculation for more than a decade (**4**). Oddly enough, in spite of the great interest in this phenomenon, a rather obvious empirical question has not been answered. There is not just one experimental variable that distinguishes the partial reinforcement schedule from the continuous reinforcement schedule but actually two. In a partial reinforcement schedule in which food is used as the reinforcing agent, the *S* gets food on some trials and no food on other trials. Is the greater resistance to extinction demonstrated by partially reinforced *S*s due to the fact that they receive food reinforcements intermittently, or is it due to the fact that they receive nonreinforcements intermittently? Or is it due to a combination of these two independent factors?

The present study sought to answer the question as to the relative influence of intermittent reinforcements and intermittent nonreinforcements on resistance to extinction in a conventional instrumental runway situation. The experimental design involved three basic groups that were hungry and thirsty during training but only hungry during extinction. Group 100 and Group 50 represented conventional, continuously and partially reinforced *S*s: Group 100 received food on every trial; Group 50 received food on half of the trials and no food on the remaining half. Group 50-50 was the key group: it received food on half of its trials and water on the other half. If the greater resistance to extinction produced by partial reinforcement is due to intermittent *food reinforcements*, then Groups 50 and 50-50 should extinguish at the same rate since both of these groups had precisely the same distribution of food reinforced trials. If, however, the greater resistance to extinction produced by the partial reinforcement schedule was due to the intermittent nonreinforcements, then Group 50 should extinguish at a slower rate than Group 50-50 since every trial of this latter group was reinforced either by food or water.

Three experiments were conducted. All three were essentially similar in design except that in the third experiment an additional 50-50 group was introduced in order to evaluate a specific methodological technique used during the extinction series of the initial two experiments.

Experiments I and II

Method

Subjects.—The *S*s were 70 experimentally naive, male albino rats 60–75 days of age purchased from a local breeder. In Exp. I there were 28 *S*s while the remaining 42 were used in Exp. II.

Apparatus.—The apparatus was a small straight runway consisting of three compartments. Two guillotine doors separated the runway into a starting box, alley, and end box. At the start of the end box was a 2-in.-high hurdle that *S* was required to surmount. A

partition projected out for a distance of 2 in. at a right angle (4 in. from the end) from one of the walls of the end box. Reinforcement was always placed behind this partition which served to prevent *S* from detecting the presence or absence of reinforcement before entering the end box.

The apparatus was painted gray and covered with a ½-in. hardware cloth. The runway was 4½ in. wide with its walls 3¾ in. high. Illumination was provided by two 15-w. lamps, one in back of the starting box and one over the door separating the alley from the end box. The timing mechanism was initiated by the raising of the door at the starting box and was stopped by *S* entering the end box.

Procedure.—The procedure for both experiments was essentially the same, consisting of three stages (pretraining, training, and testing). The pretraining stage covered six days during which *Ss* had daily 45-min. feeding and watering periods. In an attempt to make the hunger and thirst drives of approximately equal strengths, *S* was given a 2.5 gm. piece of food immediately after the food and water were removed. This drinking and eating schedule persisted throughout the pretraining and training series. During the pretraining stage *S* had experience in eating the food pellets that would later serve as reward, and in drinking from the small container which was to be used for the water rewarded trials. On the last day of pretraining each *S* was permitted to roam around the runway with the doors raised for a period of 5 min.

The three experimental groups during the training period were as follows:

Group 50 (9 *Ss* in Exp. I, 15 *Ss* in Exp. II) received food reinforcements on half of their training trials and nothing on the remaining trials. On the first trial, 5 food pellets were placed behind the partition in the end box, and *S* was permitted to eat 3. On the second trial, 3 pellets were placed in the end box, and *S* was permitted to eat 2. On all other reinforced trials, one food pellet was used. (This procedure of food reinforcements was used for all groups.) The pellets weighed .065 gm. but were humidified after which they weighed .070 gm. This procedure was followed so that the consumption of food would not so intensify the thirst drive as to interfere with food consumption on the successive reinforced trials of that day.

Group 100 (9 *Ss* in Exp. I, 13 *Ss* in Exp. II) was trained under the same conditions as Group 50 except that Group 100 received a food reinforcement on every training trial.

Group 50-50 (10 *Ss* in Exp. I, 14 *Ss* in Exp. II) received either a food reinforcement (dry pellet weighing .065 gm.) or a water reinforcement (drinking period of approximately 4 sec.) on each training trial.

There were four successive days of 10 training trials each. The sequence for each day was: Day 1, FFNFNFNNFN; Day 2, FNNFNFFNFN; Day 3, FNNFFFNNFN; Day 4, FNFFNNNFFN. F represents a food reinforced trial, while N represents a nonreinforced trial for Group 50. Group 100 had food reinforcements on both F and N trials. Group 50-50 had food reinforcements on F trials and water reinforcements on N trials.

A typical training trial proceeded as follows: 2–3 sec. after being placed in the starting box, the door was raised. When *S* entered the end box, its door was lowered. If food was present, *S* was permitted to consume it. If water was present, *S* was allowed to drink for approximately 4 sec. If it was a nonreinforced trial, *S* remained in the end box for approximately 8 sec. A period of 20 sec. elapsed between successive trials. The *S* spent this 20-sec. period in an unpainted detention box the dimensions of which were 8 × 8 × 12 in.

The extinction trials occurred on the day following the fourth day of training. On the last day of the training period water and food were removed as usual from the feeding cage. Twenty-five minutes before the extinction session *S* was given water so that during the extinction series *S* was hungry but not thirsty. Water was present in the detention box during the extinction trials to maintain the water satiated condition. Except for the absence of reinforcement, the extinction trials proceeded in the same manner as did the training trials. The criterion of extinction was that *S* failed to enter the goal box in less than 15 sec. for three successive trials. If *S* did not enter the end box within this 15 sec. time limit, he was given an additional 165 sec. to do so before he was removed. This procedure was followed because preliminary experimentation revealed that the performance of some *Ss* was seriously disrupted if they were removed from the runway instead of from the end box as they had been during training. The initial trial during the extinction series was rewarded for all groups.

Results

The results of the training trials were similar for the three groups in both experiments. On the last day of the training series there was no significant difference in the median time scores of the three groups. The medians ranged from 1 sec. to 3.5 sec.

TABLE 1

NUMBER OF RESPONSES TO REACH CRITERION OF EXTINCTION

Exp.	Group	Mean	Mdn.	σ	Range
I	100	12.4	12.0	6.4	3–28
	50	38.0	39.0	15.7	14–59
	50–50	26.6	17.5	25.0	6–68
II	100	16.7	13.0	8.4	6–33
	50	35.7	25.0	27.4	8–100
	50–50	20.4	13.5	23.4	2–100

in Exp. I, and from 2 sec. to 4 sec. in Exp. II.

Table 1 reports the number of trials required by the various groups to reach the criterion of extinction. Two *S*s in Exp. II had not reached the criterion of extinction by 100 trials and therefore were assigned extinction scores of 100. Because of the skewness of most of the distributions and the unequal variances between groups in the same experiment, the reliability of differences was estimated by the *U* test (**6**). The probabilities obtained were doubled so that the comparisons reported in Table 2 are consistent with a two-tailed test.

Table 2 shows that in both experiments partial reinforcement extinction effects were obtained, i.e., Group 50 required significantly more trials to reach the criterion of extinction than Group 100. Group 50-50, on the other hand, did not differ significantly from Group 100 in either experiment. The results suggest that a partial reinforcement procedure (Group 50) produces a greater resistance to extinction than a 50-50 procedure. In Exp. I the *P* value of the difference between Group 50 and Group 50-50 was only .19. In Exp. II the difference was significant at the .01 level. The probability of obtaining one significant difference at the .01 level in two comparisons is .02. Another way of combining these two probabilities would be to determine the probability, in two comparisons, of obtaining two differences at or below the .19 level. The answer to this question would be $.19^2$ or .04. Neither of these two methods takes into consideration that in both experiments the direction of the difference was the same (Group 50 was superior to Group 50-50). If direction is considered, the levels of significance just quoted would be halved.

TABLE 2

LEVELS OF SIGNIFICANCE OF DIFFERENCES BETWEEN EXPERIMENTAL GROUPS

Exp.	Groups 50 and 100	Groups 50 and 50–50	Groups 50–50 and 100
I	.01	.19	.41
II	.01	.01	.46

EXPERIMENT III

The results of Exp. I and II suggest that the potent variable in determining partial reinforcement extinction effect is the intermittent nonreinforced training trials rather than the intermittent reinforced training trials. The data suggested a significant difference between the performance of Group 50, as compared with those of Groups 100 and 50-50, but no significant difference between the latter two groups.

The purpose of Exp. III was to collect additional data about the phenomena reported in Exp. I and II and to investigate the possibility that the inferior performance of Group 50-50, as compared with Group 50, may have been due to the procedure used during the extinction trials. It might be argued that during extinction the end-box situation was completely new for Group 50-50, whereas the end-box situation was physically identical with that

which prevailed for Group 50 on half its training trials. Group 50, on half of their training trials, found the end box empty which was, of course, the condition that existed during extinction. Group 50-50, during extinction, not only suffered the loss of food reinforcement but also the loss of the water vessel and the water. Is it possible that the inferiority of Group 50-50 to Group 50 was due to a stimulus-generalization decrement produced by the removal of water and the water vessel during extinction? Would the resistance to extinction of Group 50-50 be equal to Group 50 if Group 50-50 had the water and/or the water vessel present during the extinction runs?

Method

Subjects.—The *S*s were 72 naive, male albino rats purchased from a local breeder. The *S*s weighed 100 to 150 gm. Upon arrival at the laboratory, they were maintained on an ad libitum food and water schedule for a period of four weeks during which time they were frequently handled by *E*.

Apparatus.—It was thought that a longer runway might provide a more sensitive measure of resistance to extinction, and therefore an apparatus different from that used in the first two experiments was used. This plan backfired somewhat, but the results did provide (with the adoption of a new response measure) information that could be used to answer the stated questions.

The apparatus was in the shape of an L with the stem being 58¾ in. in length and the arm being 13¾ in. long. The starting box was 10 in. long with a guillotine door at its end, followed by a 39½-in. alley also terminating with a guillotine door in front of the end box; 4½ in. past this latter door a 6-in. response plate was placed at a right angle to the stem. Between the response plate and the terminal wall of the end box was space for a 2½-in. glass furniture caster.

The apparatus was constructed of ¾-in. plywood, painted a light gray and covered throughout with ½-in. hardware cloth. Illumination was provided by a single 75-w. frosted light bulb suspended above a point 10 in. beyond the starting-box door. The width of the alley, as well as its height, was 4½ in.

When the starting-box door was raised ¼ in., a microswitch started an electric timer, which was stopped when *S* exerted a minimum pressure of 50 gm. on the response plate.

The detention box, used for the intertrial interval, measured 9 × 9 in. and was 9¼ in. high. It was constructed of unpainted ¼-in. plywood and was located in a dim corner of the room.

Procedure.—The procedure for Exp. III was similar to the first two except for the following modifications: (*a*) The *S*s were given 15 min. exploration sessions in which the starting-box door was closed as soon as *S* entered the alley. (*b*) On the first training day, *S* received 6 trials with the sequence of FNFNNF. The next two days had the same sequence of training trials that prevailed in Exp. I and II. On Day 4, the sequence of trials was FNNFNFNNFF; while on Day 5, the sequence was FNFFNNFNFN. (*c*) On the first food trial, *S* encountered 10 pellets in the end box and was allowed to eat 3. On the second food trial, *S* was allowed to eat 2 of the 5 pellets he found. Thereafter only one pellet was found. The intertrial interval during training varied between 30 and 60 sec. (*d*) One-half hour before the extinction trials on the day following the last day of training, *S* was given a dish of water and a water bottle in his home cage. A minimum of 20 and a maximum of 50 extinction trials were run. The criterion of extinction was three consecutive failures to depress the response plate within 60 sec. after the starting-box door was raised. The intertrial interval varied from 30 to 45 sec. The first extinction trial was not reinforced.

The major difference between this experiment and the previous two consisted in the treatment of the experimental groups. Food-and-water reward was given in identical glass dishes. Group 100 found food in their glass dish on every training trial; Group 50 found food in the glass dish on 50% of their trials while on the remaining trials the glass dish was present and empty. Group 50-50 during training received food on half of their training trials and water on the other half. During extinction, Group 50-50 was divided into two groups. Group 50-50E ran to an empty dish as did Group 50 and Group 100. Group 50-50W, during extinction, had the glass dish half filled with water. If the inferior performance of Group 50-50 in the initial two experiments was due to a stimulus-generalization decrement, then a comparison between Groups 50-50E and 50-50W would indicate whether the decrement effect was due to the removal of the water vessel or the water or both. Each of the four experimental groups had 18 *S*s.

Results

The median running times for the various groups during the terminal portion of the training series were similar. None of the four groups differed significantly from any of the others in terms of their median running time on the last five training trials. The range of these medians was from 2.7 sec. to 4.3 sec.

Table 3 reports the number of responses made by the four groups to reach the criterion of extinction (three successive failures to respond within 60 sec.). A relatively large percentage of each group failed to reach this criterion during the 50-trial extinction series. Probably as a result of this "ceiling effect," the well established, partial reinforcement extinction effect was not reliably demonstrated. The difference between Groups 50 and 100 (two-tail U test) was not significant ($P = .24$). None of the other differences was significant at the .05 level.

In spite of these negative findings, it was felt the data of Exp. III might yield results that could answer the experimental questions it was designed to test. The fact that the partial reinforcement extinction phenomenon has such a solid empirical base would seem to justify a selection of a response measure other than that originally planned. The reasoning was that if a response measure could be found that would reflect the traditional difference between a partial and a continuous reinforcement schedule, then that measure could be used to evaluate the other differences this experiment was designed to test.

TABLE 3

NUMBER OF RESPONSES TO REACH CRITERION OF EXTINCTION (60 SEC.) IN EXP. III

Group	Mean	Mdn.	σ	Range	% of *Ss* with Scores of 50
50	39.83	48.5	12.53	15–50	50
100	34.72	36	13.77	13–50	33
50-50E	36.28	45.5	15.48	9–50	39
50-50W	32.28	32.5	15.54	7–50	33

TABLE 4

NUMBER OF RESPONSES TO REACH CRITERION OF EXTINCTION (30 SEC.) IN EXP. III

Group	Mean	Mdn.	σ	Range	% of *Ss* with Scores of 50
50	31.94	28	14.26	13–50	33
100	22.11	19	12.43	7–50	11
50-50E	25.28	20	16.79	4–50	17
50-50W	22.11	16	16.70	0–50	17

It was quite easy to discover such a new response measure. It required, as one would guess, a more stringent response measure. A criterion of three successive failures to respond within 30 sec. produced a significant difference (two-tail) between Groups 50 and 100 at the .05 level.

Table 4 reports the result with this more stringent extinction criterion. There was no significant difference between the results of Group 50-50E and Group 50-50W ($P = .57$). They were combined in order to obtain a better estimate of the differences between the 50-50 schedule and the other two schedules (50 and 100). Group 50 was significantly superior ($P = .04$) to the combined Group 50-50, while there was no significant difference between the combined Group 50-50 and Group 100 ($P = .74$). These differences were consistent with those found in the initial two experiments.

DISCUSSION

The results of all three experiments are generally consistent in suggesting that partial reinforcement extinction

effects are due primarily to intermittent nonreinforced training trials rather than intermittent reinforced training trials. Whereas the intermittently nonreinforced group (Group 50) required significantly more trials to extinguish than the group that received intermittent reinforcements (Group 50-50), there was apparently no significant difference between this latter group and the group that received continuous reinforcements (Group 100). The results of Exp. III also suggested that the relative inferiority of the intermittent reinforcement group (Group 50-50) to the intermittent nonreinforcement group (Group 50) was *not* due to the greater stimulus-generalization decrement suffered by Group 50-50 in the end-box situation during extinction.

It would not be safe to conclude from our results that the effects of a 50-50 reinforcement schedule is equivalent to a continuous reinforcement schedule (Group 100). It should be noted that the variability of the 50-50 groups was significantly greater than Group 100 in Exp. I and II in which no marked ceiling effect on extinction scores operated. In this respect it would seem that the behavior of Group 50-50 is more similar to that of Group 50.

But if attention is restricted to the average resistance of the various experimental groups to extinction, one is forced to conclude that an understanding of the partial reinforcement extinction effects lies in the direction of understanding the effects of the intermittent nonreinforcements.

Although the present study has up to now been described largely in terms of empirical variables, it was not conducted in a theoretical vacuum. The results are inconsistent with any view which assumes that the greater resistance to extinction demonstrated by partially reinforced *S*s is due to their having learned to "expect," for example, food reinforcement on only some trials. If such an expectancy hypothesis were true, no difference should have occurred between the extinction rates of Group 50 and Group 50-50. Since their training involved the same schedule of intermittent food reinforcements, they should have developed similar expectations about the occurrence of food reinforcements.[1]

The problem then is to explain how the nonreinforced training trials operate to produce greater resistance during extinction. One possibility was offered by Hull's stimulus-trace hypothesis (**3**), supported by Sheffield's findings (**7**), that when training trials were distributed, the partial reinforcement extinction effect disappeared. Recently, however, Hull's stimulus-trace hypothesis has been seriously discredited. Weinstock (**9**) obtained definite evidence of a greater resistance to extinction by partially reinforced *S*s when the intertrial interval was 24 hr. Wilson, Weiss, and Amsel (**10**) also report that partial reinforcement groups were more resistant to extinction than were continuous reinforcement groups regardless of the intertrial interval during acquisition. Other experiments (**2, 8**), varying conditions other than intertrial interval, have also produced results inconsistent with Hull's hypothesis.

The experiments just cited rule out the intertrial time interval following a nonreinforced trial as the key to the understanding of the partial reinforcement extinction effect. The present experiment similarly rules out intermittent reinforced trials. Where then is the empirical locus of the partial reinforcement extinction effect? It would seem to lie in the nonreinforced training trial itself.

[1] The adequacy of this analysis depends upon *S*s in the 50-50 group having discriminated between their hunger and thirst drives. If no such discrimination had taken place, then it might be argued that the 50-50 group had essentially the same training procedure as did Group 100. Such an argument, however, would be inconsistent with the difference in the variability of extinction scores between Groups 100 and 50-50. It would also go contrary to the evidence (**1, 5**) that rats are capable of discriminating between simultaneous hunger and thirst drives.

During training the end-box responses of the partially reinforced group (Group 50) differ on reinforced and nonreinforced trials. On reinforced trials, as was the case in this experiment, the dominant end-box response is the consumption of food. This response of course cannot occur on the nonreinforced trials. Some response, however, does occur, and it is labeled, in a neutral fashion, a nonconsummatory response. It is assumed that these distinctive, nonconsummatory, end-box responses (or what some would prefer calling frustrated responses) can be conditioned, in a classical manner, to the stimuli in the end box and other parts of the apparatus. The cues produced by these anticipatory nonconsummatory responses would be, in turn, conditioned to the intrumental response.

The explanatory problem then becomes how these anticipatory nonconsummatory responses help the partially reinforced group resist extinction effects. During extinction the nonconsummatory end-box responses occur for all groups. During the early stages of extinction they would have the opportunity to work backwards (with the help of secondary reinforcement) to the earlier portions of the runway. These anticipatory responses would interfere with the anticipatory consummatory responses learned by Group 100 and Group 50-50 during the training period, but would not interfere with the anticipatory, *non*consummatory, end-box responses learned by Group 50 during their training period. As a result, an important set of response-produced cues that had been conditioned to the instrumental response would be disrupted for Group 100 and Group 50-50. Such would not be the case for Group 50 because of their nonconsummatory end-box responses learned during their training period.

It should be emphasized that the mechanisms invoked to generate this explanation are hypothetical in that the anticipatory end-box response and the cues resulting therefrom are inferred and not directly observed. Classical conditioning has been used as the model for this theorizing.

It is interesting to note that recently Wilson, Weiss, and Amsel (**10**) have also presented an analysis of the partial reinforcement extinction effect based on the concept of the fractional anticipatory goal responses. Weinstock (**9**) has also emphasized the importance of end-box responses in dealing with the partial reinforcement extinction phenomenon.

One other point should be mentioned. Reference has already been made to the greater variability during extinction produced by the partial reinforcement schedule as contrasted with the continuous reinforcement schedule. This has been noted by other investigators (**4**). The implication that the 50-50 procedure is more similar to the partial reinforcement schedule in respect to variability suggests that variability during extinction may be related to the *number* of different end-box responses occurring during training. With one end-box response, the rate of extinction of an individual *S* may be directly related to how strong the cues from its anticipatory response have been conditioned to the instrumental response. With two separate end-box responses, the speed of extinction will be a function of how strongly the distinctive cues from each anticipatory response have been conditioned to the instrumental response. The absolute and relative strengths of the associations between the response-produced cues and instrumental response may vary very markedly from *S* to *S*, and hence the variability would be much greater than the case in which the *S*'s behavior was determined by the strength of only one such S-R association. Presumably a training technique involving three different end-box responses (e.g., food, water, and nonconsummatory) would produce greater variability than one involving only two end-box responses (e.g., food and nonconsummatory).

Summary

The purpose of this study was to determine which of two factors produced the greater resistance to extinction demonstrated by partially reinforced *S*s as compared to continuously reinforced *S*s. In a partial reinforcement

schedule in which food is used as the reinforcing agent, *Ss* get food on some trials and no food on other trials. Is their greater resistance due to the fact that they receive food reinforcements intermittently, or is it due to the fact that they receive nonreinforcements intermittently?

The answer to this question was sought in the results of three experiments involving the same experimental design and employing a simple runway situation. During training the experimental groups were hungry and thirsty. Group 100 received food reinforcements on all training trials, while Group 50 had only half of their trials similarly reinforced. Group 50-50 was reinforced with food on half of their trials and water on the other half. During extinction *Ss* in all groups were hungry but water satiated. Because Group 50 and Group 50-50 had the same pattern of intermittent food reinforcements, it would be expected that, if intermittent reinforcements are responsible for partial reinforcement extinction effects, these two groups should extinguish at the same rate. Such was not the case. Group 50 required more trials to extinguish than Group 50-50, while there was no significant difference between the extinction rate of Group 50-50 and Group 100. The results suggest that the locus of the partial reinforcement extinction effects lies in the understanding of the effects of the nonreinforced training trials. An explanation was offered that involved the fractional anticipatory goal response concept.

References

1. Christie, R. The role of drive discrimination in learning under irrelevant motivation. *J. exp. Psychol.*, 1951, **42,** 13–19.
2. Crum, J., Brown, W. L., & Bitterman, M. E. The effect of partial and delayed reinforcement on resistance to extinction. *Amer. J. Psychol.*, 1951, **64,** 228–237.
3. Hull, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.
4. Jenkins, W. O., & Stanley, J. C. Partial reinforcement: a review and critique. *Psychol. Bull.*, 1950, **47,** 193–234.
5. Kendler, H. H. The influence of simultaneous hunger and thirst drives upon the learning of two opposed spatial responses of the white rat. *J. exp. Psychol.*, 1946, **36,** 212–220.
6. Mann, H. B., & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. math. Stat.*, 1947, **18,** 50–60.
7. Sheffield, J. F. Extinction as a function of partial reinforcement and distribution of practice. *J. exp. Psychol.*, 1949, **39,** 511–526.
8. Tyler, D. W., Wortz, E. C., & Bitterman, M. E. The effect of random and alternating partial reinforcement on resistance to extinction in the rat. *Amer. J. Psychol.*, 1953, **66,** 57–65.
9. Weinstock, D. Resistance to extinction of a running response following partial reinforcement under widely spaced trials. *J. comp. physiol. Psychol.*, 1954, **47,** 318–322.
10. Wilson, W., Weiss, E. J., & Amsel, A. Two tests of the Sheffield hypothesis concerning resistance to extinction, partial reinforcement, and distribution of practice. *J. exp. Psychol.*, 1955, **50,** 51–60.

(Received April 9, 1956)

Journal of Experimental Psychology
Vol. 53, No. 4, 1957

EFFECTS OF FOREPERIOD, INDUCED MUSCULAR TENSION, AND STIMULUS REGULARITY ON SIMPLE REACTION TIME [1]

WARREN H. TEICHNER

Quartermaster Research and Development Center, United States Army

Davis (**3**) has shown that the tension of the responding muscle system, as measured with surface electromyograms, increases during the foreperiod of reaction and reaches a maximum at the end of the foreperiod. He has also shown that the greater the tension level reached, the shorter the reaction time (RT). It might be deduced from this that the longer the foreperiod, the greater the tension that can be developed, and therefore, that there is an inverse relationship between length of foreperiod and RT. However, under the conditions of Davis' study this is not true, as Davis has shown, at least when foreperiods vary irregularly in length during the reaction series. Under these conditions, Davis found that maximum tension and shortest RT's were obtained with that foreperiod whose length was about average for the irregular series.

Other information concerning the relations among foreperiod, tension, and reaction time comes from studies of induced muscular tension. These studies (**1, 5, 6, 7, 8**) also report an inverse relationship between tension level and RT. However, in studies done by Freeman (**6**) and Freeman and Kendall (**7**) with irregularly presented foreperiods and induced tension levels, it was found that the foreperiod which was most effective in producing a short RT was longer with the greater than it was with the lesser of the two induced tensions used. Thus, there appears to be some support for the hypothesis that RT is inversely related to foreperiod length and that this relationship depends on the inverse relationship between tension level and RT.

On the other hand, this result appears in direct contradiction to that obtained by Davis. Since in Davis' experiment only the foreperiod was variable, whereas in Freeman's studies both foreperiod and tension magnitude were variable, it is possible that differences in stimulus regularity between the experiments may account for the difference in results. The present investigation was designed to study this problem and, in addition, to provide further information on the relationships among RT, foreperiod, and magnitude of induced tension.

The apparatus is shown in Fig. 1.[2] Muscular tension (load) was induced by having *S* pull against a calibrated spring during the entire foreperiod. Force of pull was monitored for *S* by a red, jewelled light which flashed on and off if his pulling force fluctuated. This light also initiated the foreperiod. The *S* held his arm in the same chinning position under all load conditions. The RT was for the initiation of a slight extra pull or jerk against the spring in response to the onset of a second, green, jewelled light the duration of which was 2 sec. Foreperiods and stimulus durations were controlled with Hunter timers. The RT was measured as the initial displacement of a pendulum in response to *S*'s reaction movement. This was recorded to an accuracy of .0001 sec. with a Hewlitt-Packard counter. Intertrial interval

[1] The major results of this study were presented at the 1956 meeting of the Eastern Psychological Association.

[2] The apparatus was designed and constructed by Dr. John L. Kobrick and is described elsewhere (**9**).

was monitored manually with the aid of a Standard Electric timer measuring in .01 sec.

Experiment I

Design

Experiment I used 75 soldiers as *Ss*. A randomized block design was employed in which each subgroup (cell) performed under one regularly presented combination of load and foreperiod and with either a 10- or 30-sec. intertrial interval. The *Ss* were given 24 trials per day for two successive days. Foreperiods of 2, 6, and 10 sec. were combined with loads of 5, 20, and 35 lb. except that no 35-lb. loads were used with massed practice. Although the total *N* of the experiment is fairly large, the experimental design allowed only five *Ss* per cell, and thus the reliability of the cells is somewhat lower than might be desired. Actually, Exp. I was designed as a preliminary experiment intended to give *Ss* practice and aid in the selection of foreperiod and load values for the remaining experiments. However, the results are suggestive and appear to warrant presentation in spite of their largely preliminary nature.

Results

All RT's were transformed to reciprocals in both this and subsequent experiments in order to reduce the skew commonly associated with latency measures. The reciprocal of a latency, by its very nature, provides an index of speed and shall be referred to as such. The mean reciprocals for the various conditions of the experiment are presented in Table 1. The values shown are based upon summations over all 48 trials of both experimental days since statistical analysis showed no effect due to either trials or days. Each value in this table, therefore, is based upon 240 observations.

TABLE 1

Mean Reciprocal RT for Constantly Presented Load and Foreperiod

Load (Lb.)	Foreperiod (Sec.)		
	2	6	10
30-Sec. Intertrial Interval			
5	3.28	3.10	3.06
20	3.27	2.96	3.42
35	4.09	3.17	3.30
10-Sec. Intertrial Interval			
5	3.19	3.11	2.70
20	3.20	3.99	3.02

The results shown in Table 1 with the more distributed practice condition are very difficult to interpret. The most serious problem, from the point of view of the present study, is the failure of these results to show a consistent increase in speed with increased tension. On the other hand, it may be seen that a consistent effect was obtained with the 10-sec. intertrial interval. In addition, it may be seen that the 6-sec. foreperiod was optimum with the massed trials, but no consistent optimum appeared with the distributed trials. In the latter case, the 2-sec. foreperiod was most frequently optimum.

Table 2 presents a summary of an analysis of variance of the data upon which Table 1 is based. To perform this analysis the 35-lb. row of the 30-sec. intertrial interval condition was omitted. Separate analyses of the two parts of Table 1 did not provide any different information. Inspection of Table 2 indicates that foreperiod and load were significant main effects. The simple interactions of each of these with intertrial interval was also significant as was the corresponding second-order interaction. No other effect had significance within the usual probability criteria. Thus, it would appear that differences in speed of reaction were produced by both the foreperiods and the loads used, but that these differences were not homogeneous for the two intertrial intervals. In view of these results and the fact that the 10-sec.

TABLE 2

ANALYSIS OF VARIANCE, EXP. I

Source	*df*	*MS*	*F*
Days (D)	1	4.24	—
Trials (T)	23	4.89	1.12
Loads (L)	1	39.02	8.95*
Foreperiods (F)	2	14.71	3.37**
Intertrial interval (I)	1	.10	—
D × T	23	2.34	—
D × L	1	5.13	1.18
D × F	2	6.98	1.60
D × I	1	.16	—
T × L	23	4.71	1.08
T × F	46	3.85	—
T × I	23	5.78	1.32
L × F	2	9.80	2.25
L × I	1	18.80	4.31**
F × I	2	47.79	10.96*
D × T × L	23	3.61	—
D × T × F	46	4.04	—
D × T × I	23	2.71	—
D × L × F	2	6.02	1.38
D × L × I	1	4.73	1.08
D × F × I	2	6.55	1.50
T × L × F	46	3.10	—
T × L × I	23	5.73	1.31
T × F × I	46	3.68	—
L × F × I	2	19.61	4.50**
Error	2467	4.36	n.s.***
Total	2879		

* $P < .01$.
** $P < .05$.
*** None of the pooled interactions was significant.

intertrial interval appeared to be more sensitive to the effects of induced tension, the 10-sec. interval was used in the following three experiments. The same range of loads and approximately the same range of foreperiods were used since these had been shown to be effective here.

EXPERIMENT II

Design

Experiment II was a split plot in which foreperiods of 2, 5, 8, and 11 sec. were presented 10 times each in random order making a total of 40 consecutive reactions elicited from each *S*. Forty-eight *Ss* of the previous experiment were assigned at random to subgroups of 12 *Ss*, each of which performed with a constant load of 5, 15, 25, or 35 lb. and the irregular foreperiods noted.

Results

Analysis of variance (Table 3) indicated that, of the within-*S* effects, both the differences between foreperiods and between trials were significant. The foreperiods-trials interaction was also significant. The effects of load were not significant; but since this was the confounded part of the split plot, no conclusion is warranted (**2**).

Inspection of the reciprocals on a trial-by-trial basis showed that the significance attributable to trials was due to a warm-up phenomenon which was generally overcome by the second trial. The foreperiod-trials interaction was due to the difference in the magnitude of this warm-up effect. Otherwise no trend toward either increased or decreased reaction speeds was apparent among any of the foreperiod conditions. The curves obtained were essentially parallel lines. For this reason the reciprocals were summed over trials and means obtained in order to determine the general effects of the foreperiods under

TABLE 3

ANALYSIS OF VARIANCE, EXP. II

Source	*df*	Mean Square	*F*
Between *Ss*	47	7.20	
Loads (L)	3	1.39	
Error (b)	44	7.60	
Within-*Ss*	1872	.30	
1. Foreperiods (F)	3	6.74	13.22*
2. Trials (T)	9	2.57	8.56*
3. F × T	27	1.04	4.38*
4. F × L	9	.20	
5. T × L	27	.19	
6. F × T × L	81	.20	
Error Within-*Ss*	1716	.27	
Error—1, 4	132	.50	
Error—2, 5	396	.30	
Error—3, 6	1188	.24	
Total	1919		

* $P < .01$.

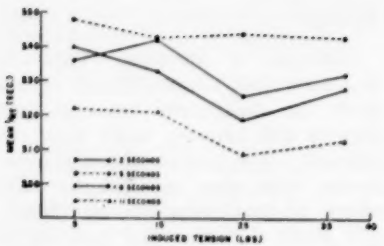

FIG. 1. Speed of reaction with different foreperiods as a function of induced tension when foreperiods were irregularly presented and loads were regularly presented during the reaction series.

the conditions of this experiment. The result of this procedure is shown in Fig. 1. Each point on this graph is based upon 120 observations. Inspection of this figure shows that the 5-sec. foreperiod was the most effective at all load conditions. This figure also suggests that the greatest speeds were achieved with the lighter loads and, in addition, suggests a possible worst load at the 25-lb. level. However, as noted above, the significance of these effects cannot be determined.

EXPERIMENT III

Design

Experiment III was also a split-plot design. The same 48 *Ss* used in the previous experiment were randomly assigned to new subgroups of 12, each of which responded with a constant foreperiod of 2, 5, 8, or 11 sec. All *Ss* were presented loads of 5, 15, 25, or 35 lb. ten times each in one 40-trial random series.

Results

Analysis of variance (Table 4) indicated that the main effects of the loads and of trials were the only significant effects. As in the previous experiment, the significance of trials could be accounted for by a quickly overcome warm-up effect. Other than this effect no trend toward either increasing or decreasing reaction speeds was discernible in the trial series. Inspection of Table 4 shows that the main effect of the foreperiods was short of a significant probability criterion. No conclusion may be drawn from this result by itself. However, Exp. II and III may be thought of as "complimentary" replicates and regarded in this fashion, both foreperiods and loads may be recognized as generally significant factors. This problem, i.e., the loss of power of the between-*Ss* error term in split-plot designs and the use of "complimentary" replicates, is discussed by Cochran and Cox (**2**).

TABLE 4

ANALYSIS OF VARIANCE, EXP. III

Source	*df*	Mean Square	*F*
Between-*Ss*	47	13.10	
Foreperiods (F)	3	30.21	2.54
Error	44	11.93	
Within-*Ss*	1872	.32	
1. Loads (L)	3	4.62	7.57*
2. Trials (T)	9	.96	3.43*
3. L × T	27	.19	—
4. L × F	9	.22	—
5. T × F	27	.34	1.21
6. T × F × L	81	.26	—
Error Within-*Ss*	1716	.32	
Error—1, 4	132	.61	
Error—2, 5	396	.28	
Error—3, 6	1188	.29	

* $P < .01$.

The major results are shown in Fig. 2 which presents mean reciprocal versus load for the different foreperiods. Each value in this graph is based upon 120 measurements. This figure shows that the 5-sec. foreperiod was optimum. It also suggests an increase in response speed with increasing load.

EXPERIMENT IV

Design

In Exp. IV, 15 of the *Ss* used in the previous experiment were selected at random and pre-

TABLE 5

ANALYSIS OF VARIANCE, EXP. IV

Source	*df*	Mean Square	*F*
Ss	14	18.72	
Series (S)	5	2.76	4.25*
Foreperiods (F)	3	12.98	25.96*
Loads (L)	3	.32	—
S × *Ss*	70	.65	
Ss × F	42	.50	
Ss × L	42	.42	
S × F	15	.42	1.11
S × L	15	.34	1.06
F × L	9	.37	1.03
S × *Ss* × F	210	.38	
S × *Ss* × L	210	.32	
Ss × F × L	126	.36	
S × F × L	45	.34	1.13
Ss × S × F × L	630	.30	
Total	1439		

* $P < .01$.

sented all 16 combinations of the four loads and four preparatory intervals in one complete randomly arranged series. At the end of this series without pause, a new random series was presented, and so on until six such series were completed.

Results

Analysis of variance of the reciprocals (Table 5) indicated that the main effects of foreperiod length and of series were both significant when evaluated over their respective interactions with *Ss*. Magnitude of load was not a significant main effect. None of the interactions was significant.

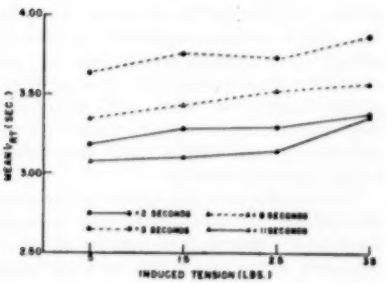

FIG. 2. Speed of reaction with different foreperiods as a function of induced tension when foreperiods were regularly presented and loads were irregularly presented during the reaction series.

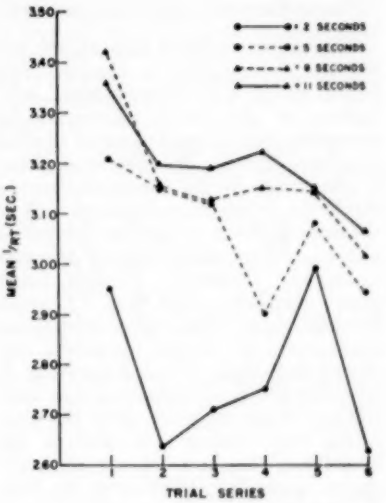

FIG. 3. Speed of reaction with different foreperiods as a function of reaction series when both foreperiod and load were variable.

The results are shown in Fig. 3 which presents the mean reciprocal versus series for each of the foreperiods. Although the behavior of the curve representing the 2-sec. foreperiod suggests a significant foreperiod series interaction as compared with the other curves, it was not significant according to the statistical analysis. In general, the effects of foreperiod are differentiated throughout the several series. Unlike the previous experiments, it appears that response speed increased the longer the foreperiod. No optimum appears. There is also a suggestion that, except for the 2-sec. interval, there was a decline in response speed from series to series.

Cross-Experiment Comparisons

The major results of Exp. II, III, and IV are summarized in Fig. 4. This figure allows a cross-experiment comparison of the mean reciprocals for foreperiod and for load summed over all other conditions. It, therefore, allows a comparison of the effects of stimulus regularity on response speed as well as the general effects of load and foreperiod under different conditions of stimulus regularity. Inspection of this figure shows that, except for Exp. IV which involved the greatest degree of stimulus irregularity, the 5-sec. foreperiod was clearly optimum. It is interesting to note that the 5-sec. foreperiod was more marked as an optimum when the foreperiod was constant than when it was irregular. The lack of an optimum shown in the foreperiod curve for Exp. IV is in accord with previous studies (**10**). However, this result extends them in suggesting the increasing relationship shown.

The results for the effects of induced tension are fairly clear. The curve obtained from Exp. III is based on the statistically most precise premises since this was a within-*S* comparison. This curve suggests that response speed increased systematically with increasing load even when the loads were presented irregularly. The results of Exp. II, which suggest a decrease in speed with load, are contradictory, but these results are based on the confounded part of the split plot. Even the curve for the irregularly presented load in Exp. IV, which was not significant, shows a slight suggestion of a rise with increasing load. The preliminary results of Exp. I were also quite clear in showing this effect with the same intertrial interval.

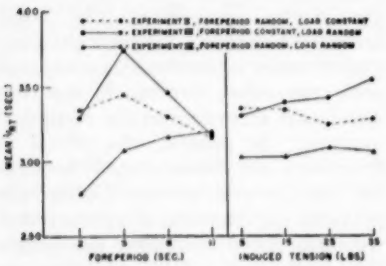

Fig. 4. Speed of reaction as a function of foreperiod and of induced tension under varying conditions of foreperiod and load regularity during the reaction series.

The probability of the stimulus condition presented to *S* is given by the product of the separate load and foreperiod probabilities. The stimulus probability of Exp. II and III was .25, and it may be seen that their general response speed level was higher than that of Exp. IV for which the stimulus probability was .06, approximately. Thus, the results indicate that stimulus irregularity is an important factor in speed of reaction.

Discussion

The results are clear in indicating that RT is inversely related to magnitude of tension. This is in agreement both with Davis' (**3**) study and other studies of induced tension (**1, 5, 6, 7, 8**). The results agree with Davis (**3**), but not with Freeman (**6**), in finding that the foreperiod length associated with the greatest speed of reaction did not change with greater tension levels. The study indicates instead that, at least with highly massed practice, there is an optimum foreperiod length which is independent of tension level in its effect on RT. In fact, since the same optimum was found in Exp. I, II and III, it would not appear that *S*s responded to the average foreperiod length, as Davis (**3**) supposed, but rather that there really is an optimum even when the foreperiod length varies randomly.

The results are also clear in showing the effects of foreperiod and tension regularity. When only one of these is irregular and its probability is at least .25, a definite foreperiod optimum appears. For the same probability, the results suggest that speed of reaction is greater when the tension level is irregular than when it is the foreperiod which is

irregular. When both are irregular and their combined probability small, speed of reaction is depressed and RT does vary inversely with foreperiod length. Herein may lie the explanation for Freeman's (6) finding longer foreperiods and shorter RT's with increased tension levels, whereas Davis (3) found maximum tension and shortest RT at what was apparently an optimum foreperiod interval. That is, in Davis' study, tension was always constant and minimum. The only irregularity was in the foreperiod. Although he used a large number of foreperiods, these differed in length by only hundredths of a second. Actually, they seem to fall most reasonably into five categories of foreperiod length which differed successively by .5 sec. Freeman (6), on the other hand, used four foreperiods (2, 4, 8, and 10 sec.) and two tension levels. Thus, in Davis' study only the foreperiod was variable, and its probability was about .20; whereas in Freeman's study both foreperiod and tension were variable, and their combined probability was .12. The present results suggest that the apparent contradiction of the two studies was the result of this difference in stimulus regularity rather than anything that might be attributed directly to the effects of tension level.

Other results of the present investigation are of considerable interest since they have implications both for studies of performance acquisition and performance decrement. The failure of any of the experiments to show important practice effects is not without precedent (9). Suggestive information relevant to this problem may be found in a recent study by Farber and Spence (4). Inspection of their Fig. 1 indicates that male *S*s generally reached maximum reaction speed by Trial 8 of a 16-trial training series. The maximum gain in mean reaction speed appears to be about 2%. However, the same figure shows that with the exception of one of the four female groups, the reaction speed of the female *S*s increased throughout the series. The maximum gain in mean reaction speed appears to be about 25%. In view of the well-known (9) difference in RT between sexes, sex differences in RT acquisition might account for failure of the present study, and others, to find important practice effects. This hypothesis at least deserves further study and should be considered as a problem for experimental control in learning studies. Along the same line it is interesting to speculate that the sex difference in RT *acquisition* is a cultural phenomenon, men having more opportunity to engage in activities requiring fast RT's than women, whereas the sex difference in *speed* of reaction is a genetic phenomenon as evidenced by the fact that men still have shorter RT's than women even after considerable practice (4).

The results of the present study may be limited to highly massed trials. Although the main effects of the intertrial interval were not significant in Exp. I, the results of that experiment suggest that the effects of foreperiod and of load may depend on the intertrial interval used. In this regard, it is important to note that a decrement in reaction speed was observed with continued performance under the conditions of Exp. IV. Since there is no evidence of learning in any of the results, it would appear that the decrement exhibited was the result of muscular fatigue, failure to maintain attention, or a decrease in motivation. In fact, the lack of important practice effects among men suggests that the RT elicited from men may be an excellent method for studying intertrial interval and performance decrement as a primarily inhibitory and/or motivational phenomenon.

Summary

Four experiments were performed with male *S*s to study the effects on RT of foreperiod length, magnitude of induced muscular tension, regularity of presentation, and intertrial interval as an attempt to resolve an apparent contradiction in results between two previous studies. In Exp. I, *S*s performed under constant, regularly presented foreperiod-load combinations and with one of two intertrial intervals. Experi-

ments II and III were split plots conducted with massed trials. In Exp. II, magnitude of tension was held constant for each subgroup, while foreperiods of different length were presented irregularly; in Exp. III the procedure was the reverse of Exp. II. In Exp. IV, both tension magnitude and foreperiod length were presented irregularly. The results appear to warrant the following conclusions:

1. Foreperiod length and magnitude of muscular tension are independent in their effects on RT.

2. RT varies inversely with magnitude of muscular tension except for combined foreperiod-tension irregularity of presentation of a high degree. In this case, magnitude of tension does not appear to affect RT.

3. Except for high degrees of foreperiod-tension irregularity or presentation, there is an optimum foreperiod of reaction. At least with massed practice for the present task, the optimum foreperiod is 5–6 sec.

4. When both foreperiod and tension level are presented irregularly and their combined probability of occurrence is low, RT varies inversely with length of foreperiod.

5. All of the above conclusions may depend on the intertrial interval employed.

6. At least under the conditions of the present study, repeated elicitation has no effect on RT except under the most irregularly presented stimulus conditions, in which case RT exhibits a decrement with continued elicitation.

References

1. Angell, F. Duration, energy, and extent of reaction movements. *Amer. J. Psychol.*, 1919, **30,** 224–236.
2. Cochran, W. G., & Cox, G. M. *Experimental designs.* New York: Wiley, 1950.
3. Davis, R. C. Set and muscular tension. *Indiana Univer. Publ.*, Sci. Ser. No. 10, 1940.
4. Farber, I. E., & Spence, K. W. Relation of stimulus intensity, practice anxiety, and sex to simple RT and temporal generalization. Iowa City, Iowa: State Univer. of Iowa, 1955. (Tech. rep. No. 4, ONR Contract N9onr-93802.)
5. Freeman, G. L. Facilitative and inhibitory effects of muscular tension on performances. *Amer. J. Psychol.*, 1933, **45,** 17–52.
6. Freeman, G. L. The optimal muscle tensions for various performances. *Amer. J. Psychol.*, 1938, **51,** 146–151.
7. Freeman, G. L., & Kendall, W. E. The effect upon reaction time of muscular tension induced at various preparatory conditions. *J. exp. Psychol.*, 1940, **27,** 136–148.
8. Knott, J. R. Some effects of mental set upon the electro-physiological processes of the human cerebral cortex. *J. exp. Psychol.*, 1939, **24,** 384–405.
9. Kobrick, J. L. A device for measuring reaction time during induced muscular tension. Quartermaster Res. and Develpm. Center, Tech. Rep. (In press.)
10. Teichner, W. H. Recent studies of simple reaction time. *Psychol. Bull.*, 1954, **51,** 128–149.
11. Woodrow, H. The measurement of attention. *Psychol. Monogr.*, 1914, **17,** No. 5 (Whole No. 76).

(Received April 11, 1956)